

**Struktur- und Funktionsanalyse der  
Protein-basierten RNase P des  
hyperthermophilen Bakteriums *Aquifex  
aeolicus* und homologer Enzyme**

Dissertation

zur

Erlangung des Doktorgrades  
der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Pharmazie der  
Philipps-Universität Marburg

vorgelegt von

**Nadine Bianca Wäber, M.Sc**

aus Lich

Marburg/Lahn 2019

Erstgutachter: **Prof. Dr. Roland Hartmann**  
Zweitgutachter: **Prof. Dr. Peter Friedhoff**

Eingereicht am: **01.11.2019**

Tag der mündlichen Prüfung am: **13.12.2019**

Hochschulkennziffer: 1180

## Zusammenfassung

Die Reifung einer transfer-RNA (tRNA) ist ein essenzieller Schritt in der Proteinbiosynthese in allen drei Domänen des Lebens. Dabei sind mehrere Schritte notwendig, um von einer Vorläufer- zu einer maturen tRNA zu gelangen. Die Spaltung der 5'-Flanke wird einzig durch die Ribonuklease P (RNase P) vermittelt. Es gibt zwei Arten der RNase P, die eine besteht aus einem mit ein bis zehn Proteinuntereinheiten assemblierten Ribozym, die andere ist eine rein Protein-basierte Form (PRORP), welche zunächst nur in Eukaryonten identifiziert wurde. Das Ribozym konnte in allen drei Domänen des Lebens nachgewiesen werden.

Im Kontext dieser Arbeit wurde eine neue Art der Protein-basierten RNase P (AaRP) in dem hyperthermophilen Bakterium *Aquifex aeolicus* charakterisiert. Im Gegensatz zu den bereits bekannten etwa 60 kDa großen Enzymen aus Pflanzen ist dieses Protein deutlich kleiner (~23 kDa) und weist erkennbar nur eine katalytische Domäne auf. Substratbindungsdomänen konnten nicht bioinformatisch nicht identifiziert werden. Das aktive Zentrum konnte durch gezielte Mutationsanalysen charakterisiert sowie die enzymatische Aktivität *in vitro* nachgewiesen werden. Zusätzlich konnte in Kooperation mit einer Wiener Arbeitsgruppe in zwei unabhängigen *in vivo* Experimenten gezeigt werden, dass AaRP in der Lage ist sowohl die endogene RNase P von *E. coli* (eine RNA plus eine Proteinuntereinheit), als auch die komplexere Form der Bäckerhefe (eine RNA plus neun Proteinuntereinheiten) funktional zu ersetzen.

Durch bioinformatische Analysen konnten Homologe von AaRP (HARP) in Bakterien und Archaeen identifiziert werden. Es wurden insgesamt jeweils fünf bakterielle und archaeale HARPs exemplarisch ausgewählt. Im Gegensatz zu *A. aeolicus* besitzen alle weiterführend untersuchten Organismen zusätzlich eine RNA-basierte Form der RNase P, welche zum Teil bereits in anderen Publikationen beschrieben wurde. Es zeigte sich, dass die HARPs der Archaeen *Methanosarcina mazei*, *Methanothermobacter thermautotrophicus*, sowie *Haloferax volcanii* ebenfalls die 5'-Enden von tRNAs prozessieren können. In Kooperation mit einer Ulmer Arbeitsgruppe konnten in *H. volcanii* Deletionsmutanten des Proteins konstruiert und weiterführend untersucht werden. Durch eine Kooperation mit einer Kieler Arbeitsgruppe konnte dies

ebenfalls in *M. mazei* durchgeführt werden. Bei den bakteriellen HARPs zeigte sich, dass die HARPs von *Thermodesulfatator indicus* und *Methylophilum infernum* ebenfalls dazu in der Lage waren, Prä-tRNAs zu prozessieren. Neben den HARPs wurden auch die RNA-basierten RNase P Enzyme auf ihre tRNA-Prozessierungs-Aktivität hin untersucht, insofern diese noch nicht in Veröffentlichungen beschrieben wurde.

Im Gegensatz zu AaRP besitzt PRORP eine *Pentatricopeptide-repeat*-Domäne, welche an der Substrat-Bindung beteiligt ist. Um mehr über die Substratspezifität von AaRP zu erfahren, wurden verschiedene Varianten einer Modell-Prä-tRNA in Spaltungsstudien untersucht. Dabei zeigte sich, dass AaRP sowohl Merkmale einer bakteriellen RNase P als auch einer eukaryotischen Protein-basierten Form aufweist. So spielt beispielsweise, ähnlich zu den bakteriellen RNase P-Enzymen, die Länge der 5'-Flanke eine Rolle bei der Substraterkennung durch AaRP; dies ist bei den PRORP-Enzymen hingegen irrelevant. Demgegenüber zeigt das 3'-Ende der Prä-tRNA keine Interaktion mit AaRP oder PRORP, welches bei den RNA-basierten bakteriellen RNase P-Enzymen essenziell ist.

Zusätzlich sollte die Struktur von AaRP und *H. volcanii* HARP aufgeklärt werden, um weitere Einblicke in die Substratbindung zu gewinnen. Während die Kernspinresonanz-Spektroskopie nicht erfolgreich war, konnte mittels Gelfiltration, dynamischer Lichtstreuung und Circular dichroismus-Spektroskopie gezeigt werden, dass AaRP fast ausschließlich  $\alpha$ -helikal ist und als großer homooligomerer Komplex (~420 kDa) vorliegt. Das *H. volcanii* HARP hingegen besitzt mehr  $\beta$ -Faltblatt Anteile und liegt als bedeutend kleinerer Komplex (~90 kDa) vor. Durch analytische Ultrazentrifugation und kryoelektronenmikroskopische Analyse konnte bestimmt werden, dass die Komplexbildung von AaRP ein heterogener Vorgang ist, da sich verschiedene Größen identifizieren ließen. Die Kristallisationsexperimente waren nur bedingt erfolgreich und die erhaltenen Kristalle waren vorwiegend klein. Eine genauere Analyse dieser Kristalle sowie eine Optimierung der Kristallisations-Bedingungen konnte im zeitlichen Rahmen der vorliegenden Arbeit nicht mehr durchgeführt werden.



## Abstract

The transfer RNA (tRNA) maturation is an essential step in the biosynthesis of proteins in all domains of life. It is a multistep process to get the precursor tRNA (pre-tRNA) into a mature form. The cleavage of the 5'-leader and the 3'-trailer is performed by different ribonucleases (RNases). Furthermore, single nucleotides are modified by a number of various enzymes. Whereas processing of the 3'-trailer can occur by diverse RNases, RNase P is the unique enzyme for cleavage of the 5'-leader. The bacterial RNase P is composed of a catalytic RNA and a small protein subunit. The RNA-based RNase P exists in all domains of life and the protein content in the ribonucleoprotein increases with the complexity of the organism. In eukaryotes a protein-based form of RNase P was identified which can be either a complex of three different proteins (e.g. in human mitochondria) or a singular protein subunit called PRORP (e.g. in *Arabidopsis thaliana*).

In the present thesis a novel single subunit RNase P enzyme was characterized in the hyperthermophilic bacterium *Aquifex aeolicus*, called AaRP. It is the smallest form known so far (~23 kDa) and consists of only a PIN-like nuclease domain. In contrast PRORP enzymes consists of three domains: (i) a pentatricopeptide-repeat (PPR) motif domain, (ii) a zinc finger binding motif domain and (iii) a NYN-metallonuclease domain. The active site of AaRP was identified by point mutation analysis. AaRP was able to replace the endogenous RNase P of *E. coli* (one RNA plus one protein-subunit) as well as the complex nuclear RNase P of yeast (one RNA plus nine protein-subunits) *in vivo*.

By bioinformatic analyses a small number of homologs of AaRP (HARP) could be identified in some bacteria and archaea. All these organisms also contain the genes for the classical RNA-based RNase P. The HARPs from three archaea (*Methanosarcina mazei*, *Methanothermobacter thermautotrophicus*, *Haloferax volcanii*) and two bacteria (*Thermodesulfatator indicus* and *Methylophilum infernorum*) were selected for further biochemical characterization. All analyzed archaeal and bacterial HARPs were able to remove the 5'-leader of pre-tRNAs by site-specific endonucleolytic cleavage. Deletion mutants of *H. volcanii* HARP, as well as *M. mazei* HARP were constructed and characterized. Further, the bacterial RNA-based RNase Ps from *T. indicus* and

*M. infernorum* were analyzed for tRNA-processing functionality *in vitro* and *in vivo*.

Substrate specificity and recognition by AaRP is still unclear since no RNA-binding motif could be identified. Therefore, biochemical studies were designed to identify the substrate determinants of AaRP/HARP. The studies showed that interaction with the 5'-leader seems to be important for tRNA recognition by AaRP/HARP since the length of the 5' flank has an influence on cleavage efficiency. This is similar to bacterial RNA-based RNase P where the protein subunit also interacts with positions in the 5' leader of the precursor tRNA. In contrast, as described for PRORP, the AaRP/HARP enzymes obviously make no interactions with extensions of the substrate 3'-end (CCA or trailer). Structure analysis could give further informations to substrate recognition. While analysis of AaRP/HARP by nuclear magnetic resonance spectroscopy was unsuccessful, dynamic light scattering experiments indicated that AaRP assembles in a large complex (~420 kDa). This finding was corroborated by gel filtration experiments. Results of a circular dichroism spectroscopy analysis indicate that AaRP has a nearly complete  $\alpha$ -helical structure. In contrast *H. volcanii* HARP contains more  $\beta$ -sheets and assembles in a smaller complex (~90 kDa). Analytical ultracentrifugation and cryo-electron microscopy showed a heterogenous distribution of AaRP even after gel filtration indicating a heterogenous complex formation. Crystallization screens (manual and automatically) yielded only small crystals. A more detailed analysis of those crystals and an optimization of the crystallization process was not able in the available time frame of the present work.

# Inhaltsverzeichnis

<b><u>Zusammenfassung</u></b> .....	<b>3</b>
<b><u>Abstract</u></b> .....	<b>5</b>
<b>Abbildungs- und Tabellenverzeichnis</b> .....	<b>10</b>
<b>1 Einleitung</b> .....	<b>11</b>
<b>1.1 Transfer RNAs in Bakterien</b> .....	<b>11</b>
<b>1.2 Die Geschichte der Ribonuklease P</b> .....	<b>13</b>
1.2.1 Die RNase P als Ribozym .....	14
1.2.2 Die RNase P als (singuläres) Protein .....	16
<b>1.3 Substraterkennung und Prozessierung durch RNase P</b> .....	<b>18</b>
1.3.1 RNA-basierte RNase P .....	19
1.3.2 Protein-basierte RNase P .....	21
<b>1.4 <i>Aquifex aeolicus</i></b> .....	<b>21</b>
<b>2 Zielsetzung</b> .....	<b>22</b>
<b>3 Material und Methoden</b> .....	<b>24</b>
<b>3.1 Bakterienstämme</b> .....	<b>24</b>
3.1.1 Lagerung und Kultivierung der Bakterienstämme .....	25
3.1.2 Herstellung chemisch-kompetenter Bakterien .....	25
3.1.3 Herstellung elektro-kompetenter Bakterien .....	26
3.1.4 Transformation von genetischem Material (Hitzeschock) .....	26
3.1.5 Transformation von genetischem Material (Elektroporation) .....	27
3.1.6 Kolonie Polymerase-Ketten-Reaktion (PCR) .....	27
<b>3.2 Klonierung von Vektoren zur Charakterisierung rekombinanter Proteine</b> .....	<b>28</b>
3.2.1 Polymerase-Ketten-Reaktion (PCR) .....	29
3.2.2 Agarose-Gelelektrophorese .....	32
3.2.3 Restriktionsverdau .....	32
3.2.4 Gelextraktion von DNA .....	33
3.2.5 Ligation .....	33
3.2.6 Plasmidisolation .....	33
3.2.7 Isolation von genomischer DNA .....	34
3.2.8 Klonierung der pET28a(+)-Vektoren .....	34
3.2.9 Klonierung der pDG148(+) Vektoren .....	35
3.2.10 Klonierung der pGEX-mod Vektoren .....	36
3.2.11 <i>In vitro</i> Mutagenese von Aq880 und Mth243 .....	39
3.2.12 Sequenzierung der Vektoren .....	42
<b>3.3 Herstellung von verschiedenen Prä-tRNA Substraten</b> .....	<b>42</b>
3.3.1 Vektoren mit diversen Varianten der Prä-tRNA <sup>Gly</sup> .....	43
3.3.2 Herstellung der Template für die T7 Transkription .....	43

3.3.3	Phenol-Chloroform-Reinigung	44
3.3.4	Ethanol-Fällung	45
3.3.5	T7 <i>in vitro</i> Transkription	45
3.3.6	Denaturierende Polyacrylamid-Gelelektrophorese (PAGE)	46
3.3.7	Reinigung der T7-Transkription mittels Gel Elution	47
3.3.8	Radioaktive Markierung der Substrate am 5'-Ende	48
<b>3.4</b>	<b><i>In vitro</i> Analysen von rekombinanten Proteinen</b>	<b>48</b>
3.4.1	Expression der rekombinanten Proteine	48
3.4.2	Proteinreinigung mittels des Batchverfahrens	49
3.4.3	Dialyse von gereinigten Proteinen	50
3.4.4	Proteinreinigung für die Kristallisation	51
3.4.5	Proteinreinigung der T7 RNA Polymerase	52
3.4.6	Bestimmung der Proteinkonzentration mittels Bradford	52
3.4.7	Aktivitätsanalyse der Proteine (nicht-radioaktiv)	53
3.4.8	Aktivitätsanalyse der Proteine (radioaktiv)	53
3.4.9	Bestimmung des $K_M$ und $V_{max}$ ( <i>single-turn-over</i> )	54
3.4.10	Bestimmung des $K_M$ und $V_{max}$ ( <i>multiple-turn-over</i> )	54
3.4.11	RNA-Bindung: DRaCALAs	55
3.4.12	Oligomerisierungsanalysen mit Glutaraldehyd	56
<b>3.5</b>	<b>Kristallisation von Proteinen</b>	<b>56</b>
3.5.1	Kristallisation im „ <i>sitting-drop</i> “ Verfahren	56
3.5.2	Kristallisation im „ <i>hanging-drop</i> “ Verfahren	57
3.5.3	Analyse der Kristalle im Röntgengerät	58
<b>3.6</b>	<b>Chromatographie</b>	<b>58</b>
3.6.1	Affinitätschromatographie	58
3.6.2	Anionenaustauschchromatographie	59
3.6.3	Kationenaustauschchromatographie	60
3.6.4	Größenausschlusschromatographie	61
<b>3.7</b>	<b><i>In vivo</i> Charakterisierung von Proteinen</b>	<b>62</b>
3.7.1	Komplementationsstudien mittels <i>E. coli</i> BW	63
3.7.2	Komplementationsstudien mittels <i>B. subtilis</i> d7	64
3.7.3	Analyse der Protein-Expression in <i>E. coli</i> BW	64
3.7.4	Analyse der Protein-Expression in <i>B. subtilis</i> d7	65
3.7.5	Analyse von Fusionsproteinen in Komplementationsstudien	65
3.7.6	SDS-PAGE	66
3.7.7	Coomassie-Färbung	66
3.7.8	Western Blot	67
3.7.9	TGX Stain-Free Gele	67
<b>4</b>	<b>Ergebnisse und Diskussion</b>	<b>68</b>
<b>4.1</b>	<b><u>Publikation 1:</u> Minimal and RNA-free RNase P in <i>Aquifex aeolicus</i></b>	<b>68</b>

4.1.1	Originalpublikation	69
4.1.2	Supplementary Informationen	76
<b>4.2</b>	<b><u>Publikation 2:</u> Homologs of <i>Aquifex aeolicus</i> protein-only RNase P are not the major RNase P activities in the archaea <i>Haloferax volcanii</i> and <i>Methanosarcina mazei</i></b>	<b>124</b>
4.2.1	Originalpublikation	126
4.2.2	Supplementary Informationen	134
<b>4.3</b>	<b>Substraterkennung der RNase P von <i>A. aeolicus</i></b>	<b>148</b>
4.3.1	Substratvariationen am 5'- oder 3'-Ende der Prä-tRNA	150
4.3.2	Minimalsubstrate	152
4.3.3	Akzeptorstamm Variationen	155
<b>4.4</b>	<b>Bakterielle Protein-basierte RNase P Enzyme</b>	<b>156</b>
<b>4.5</b>	<b>Strukturaufklärung protein-basierter RNase P</b>	<b>161</b>
4.5.1	Bestimmung des Oligomerisierungsgrads	162
4.5.2	Oligomerisierungsgrad Bestimmung durch analytische Ultrazentrifugation	164
4.5.3	Größenbestimmung mittels dynamischer Lichtstreuung	165
4.5.4	Strukturanalyse durch Aminosäureaustausche	167
4.5.5	Strukturanalyse mittels Circular dichroismus-Spektroskopie	168
4.5.6	Strukturaufklärung mittels Kernspinresonanzspektroskopie	169
4.5.7	Kristallisation der RNase P von <i>Aquifex aeolicus</i>	170
4.5.8	Kristallisation des HARPs von <i>Haloferax volcanii</i>	175
4.5.9	Diskussion der Kristallisation	176
<b>5</b>	<b>Ausblick</b>	<b>178</b>
<b>6</b>	<b>Literatur</b>	<b>180</b>
<b>Anhang</b>		<b>190</b>
	<b>BioSpektrum – Artikel 2018</b>	<b>190</b>
	<b>Artikel in „Encyclopedia of Microbiology“ – 2019</b>	<b>194</b>
	<b>Gen- und Proteinsequenzen</b>	<b>203</b>
	<b>NCBI- Annotation von Minf0924</b>	<b>208</b>
	<b>Publikationen</b>	<b>211</b>
	Veröffentlichungen in Fachzeitschriften	211
	Veröffentlichungen in Büchern	211
	Veröffentlichungen auf Fachkonferenzen	211
	<b>Lebenslauf</b>	<b>214</b>
	<b>Danksagung</b>	<b>215</b>
	<b>Erklärung</b>	<b>216</b>

## Abbildungs- und Tabellenverzeichnis

Abbildung 1: Übersicht der Prozessierung von tRNAs .....	12
Abbildung 2: Schematische und strukturelle Darstellungen RNA-basierter RNase Ps .....	16
Abbildung 3: Schematische und strukturelle Darstellungen Protein-basierter RNase P .....	18
Abbildung 4: Übersicht von Prä-tRNA-Strukturen und RNase P-Domänen .....	20
Abbildung 5: Strukturvariationen der Prä-tRNA für die Substraterkennung .....	149
Abbildung 6: Graphische Darstellung der Spaltungseffizienzen .....	151
Abbildung 7: Prozessierung von Prä-tRNA Substratvarianten .....	154
Abbildung 8: Prozessierung von Akzeptorstamm Varianten .....	157
Abbildung 9: Analyse rekombinanter bakteriellen HARPs auf RNase P Funktion .....	158
Abbildung 10: Komplementationsanalysen der bakteriellen HARPs und P Proteine .....	162
Abbildung 11: Gelfiltrations-Chromatogramm .....	164
Abbildung 12: Sedimentationsverhalten von AaRP .....	165
Abbildung 13: Größenanalyse von AaRP mittels Dynamischer Lichtstreuung .....	166
Abbildung 14: Analyse der <i>Coiled-Coil</i> Vorhersage für AaRP .....	168
Abbildung 15: CD-Analyse von AaRP und Hv-HARP .....	169
Abbildung 16: 1D <sup>1</sup> H Spektrum von AaRP .....	170
Abbildung 17: Kristalle von AaRP mit Diffraktionsanalysen .....	173
Abbildung 18: Kristalle von AaRP/HARP des MarXtal-Screens .....	174
Abbildung 19: Kristalle und Sphäroliten von Hv-HARP .....	176
Tabelle 1: Auflistung der in dieser Arbeit verwendeten Bakterienstämme .....	24
Tabelle 2: Auflistung der Oligonukleotide für die Kolonie-PCR .....	28
Tabelle 3: Thermoprotokoll für eine Kolonie-PCR .....	28
Tabelle 4: Übersicht der verwendeten Vektoren .....	30
Tabelle 5: Auflistung der extern synthetisierten Plasmide für die Klonierung nach pET28a(+) .....	35
Tabelle 6: Auflistung der Oligonukleotide für die Klonierung in pET28a(+) .....	36
Tabelle 7: Auflistung der Oligonukleotide für die Klonierung in pDG148(+) .....	37
Tabelle 8: Auflistung der Oligonukleotide für die Klonierung in pGEX-mod .....	39
Tabelle 9: Auflistung der Primer für die in vitro Mutagenese von aq_880 .....	40
Tabelle 10: Auflistung der Primer für die in vitro Mutagenese von mth_243 .....	41
Tabelle 11: Auflistung der Plasmide für 5'-Längen Varianten von Prä-tRNA <sup>Gly</sup> .....	43
Tabelle 12: Auflistung der Oligonukleotide für die Herstellung von Minimalsubstraten .....	44
Tabelle 13: Ansatz einer T7 in vitro Transkription .....	46
Tabelle 14: Übersicht der Lagerpuffer für Proteine .....	50
Tabelle 15: Übersicht der Kristallisationsbedingungen .....	57
Tabelle 16: Auflistung der Standardproteine für die Größenausschlusschromatographie .....	62
Tabelle 17: Ergebnisse der Analysen zur Substraterkennung .....	152

## 1 Einleitung

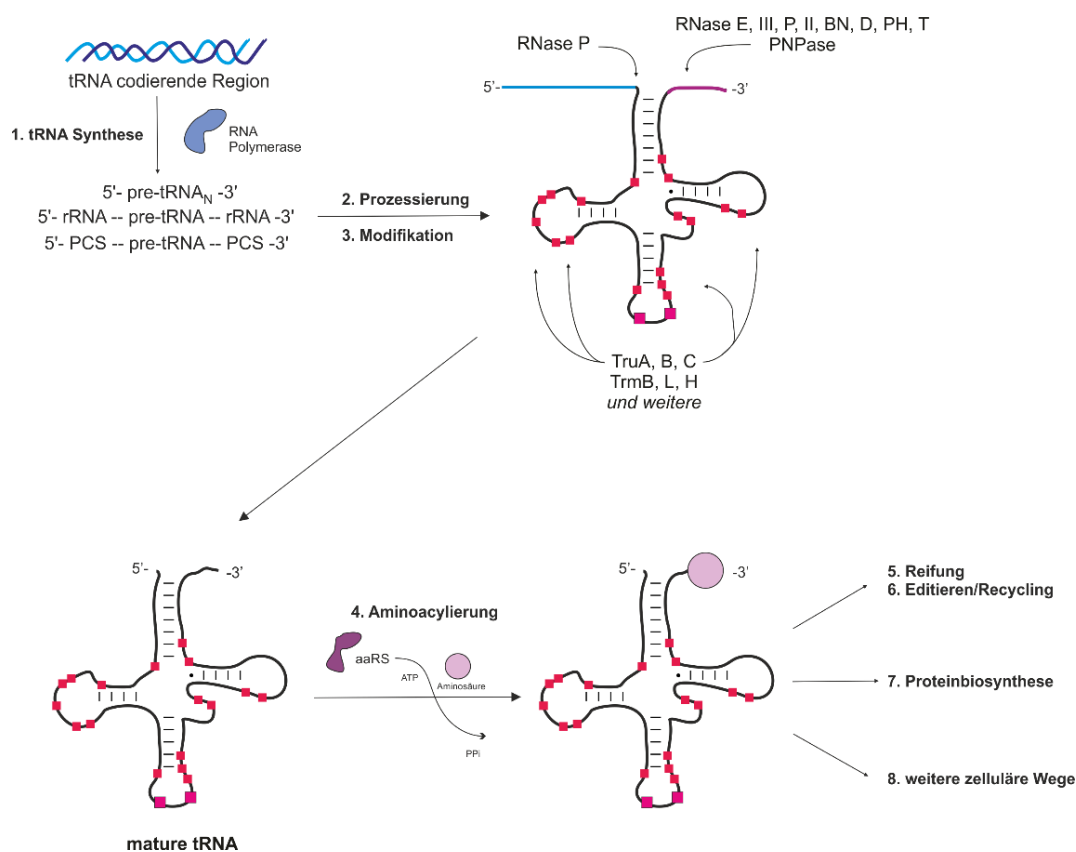
### 1.1 Transfer RNAs in Bakterien

Transfer RNAs (tRNAs) sind ein essenzieller Bestandteil der Proteinbiosynthese, welche die Aminosäuren in das Ribosom transportieren. Bis auf eine Ausnahme werden diese Moleküle als Vorläufer tRNAs (Prä-tRNAs) transkribiert, die durch mehrere Stufen prozessiert werden müssen, bevor sie in die verschiedenen zellulären Prozesse eingreifen können.

Die Prä-tRNAs können sowohl einzeln als auch polycistronisch transkribiert werden. Ebenso können die Prä-tRNAs zwischen ribosomalen RNAs (rRNA, Abbildung 1), oder Protein-codierenden Sequenzen (PCS, Abbildung 1) liegen [1]. Nach der Transkription der Prä-tRNAs durch die RNA Polymerase werden die Moleküle sowohl am 5'- und 3'-Ende durch Nukleasen prozessiert als auch durch diverse Enzyme modifiziert. Der Überhang am 5'-Ende der Prä-tRNA wird durch die RNase P entfernt, welche als einzige Nuklease dort angreift [2-4]. Die 3'-Prozessierung wird typischerweise durch eine endonukleolytische Spaltung stromabwärts des CCA-Motivs initiiert [5-7]. Dies wird meist vor der 5'-Flanken Spaltung und der finalen 3'-Prozessierung durchgeführt [8-10]. Unter anderem könne RNase E und RNase III diese Spaltungen durchführen und spielen bei polycistronischen Transkripten eine Rolle [11-13]. Das finale Entfernen der Nukleotide am 3'-Ende durch RNase D ist essenziell, denn diese Nuklease ist in der Lage das meist schon codierte CCA-Motiv intakt zu lassen [14, 15]. Weitere Nukleasen sind RNase BN, PH, T und PNPase [16]. Sobald die 5'-Flanke entfernt und am 3'-Ende das CCA-Motiv erkennbar ist, ist die tRNA bereit mit seiner spezifischen Aminosäure beladen zu werden. Bevor jedoch die Aminoacylierung durchgeführt werden kann, werden verschiedene Nukleotide der tRNA chemisch modifiziert. Zur Zeit sind mehr als 100 verschiedene Modifikationen bekannt, wobei die Positionen 34 und 37 der tRNA die meisten unterschiedlichen chemischen Veränderungen aufweisen können (■, Abbildung 1) [17]. Alle Auswirkungen, die diese Modifikationen mit sich bringen sind bisher nicht bekannt, allerdings scheinen sie eine fundamentale Rolle in der Stabilisierung, Erkennung und Steuerung von zellulären Prozessen zu spielen, denn die Modifikationen unterscheiden sich zwischen einzelnen Lebensbedingungen [18]. Die Modifikationen werden durch verschiedene

Enzyme vermittelt, wie zum Beispiel TruA, TruC oder TruD, welche eine  $\Psi$ -Modifikation anbringen [19-22]. Da in den Wildtyp-Stämmen der meisten Bakterien keine Prä-tRNAs gefunden werden können, sondern nur in Mutanten, welche defizient in der tRNA Prozessierung sind, liegt es nahe, dass die ganzen Prozesse während, oder direkt nach der Synthese stattfinden müssen [2, 18].

Der Hauptteil der aminoacylierten tRNAs gelangt mit Hilfe des Elongationsfaktors Tu (EF-Tu) zum Ribosom und wird für die Proteinbiosynthese verwendet [18]. Allerdings wird ein kleiner Teil der tRNAs auch für die Peptidoglycan-Synthese, Antibiotika-Resistenz/Synthese oder Phospholipid-Modifikationen der Membranen genutzt. Zudem können tRNAs auch in Toxin/Antitoxin-Systemen eine Rolle spielen, indem sie zum Beispiel von dem Toxin VapC in der Anticodonschleife gespalten und somit inaktiviert werden [18, 23].



**Abbildung 1: Übersicht der Prozessierung von tRNAs**

Es ist schematisch der Weg einer tRNA von der Transkription des genetischen Bereichs bis hin zu der maturen und aminoacylierten tRNA dargestellt. Dabei sind die einzelnen Schritte nummeriert, sowie ein Teil der essenziellen Enzyme für die einzelnen Prozesse aufgeführt (nach [18]).



## 1.2 Die Geschichte der Ribonuklease P

Anfang der 1970er-Jahre wurden aus Mutanten von *E. coli* Tyrosin-tRNA Moleküle isoliert, welche noch eine 5'-Flanke aufwiesen [2]. Die Spaltung dieser 5'-Flanke konnte im Anschluss *in vitro* detektiert werden [24]. Es wurde gezeigt, dass auch andere tRNAs mit 5'-Flanken akkumulieren (in thermosensitiven *E. coli* Stämmen) [25]. Erst Ende der 1970er-Jahre konnte nachgewiesen werden, dass das für die 5'-Prozessierung verantwortliche Enzym ein Ribonukleoprotein ist, bestehend aus einem Protein und einer RNA-Komponente [26]. Das Ribonukleoprotein-Enzym wurde RNase P genannt. In dieser Publikation wurde der RNA Komponente noch eine untergeordnete Rolle zugewiesen. Dabei wurde vermutet, dass die RNA zur Stabilisierung sowie Interaktion mit dem Substrat benötigt wird [26]. Um weitere Erkenntnisse über dieses neue Enzym zu bekommen, wurde die RNase P durch chromatographische Methoden in ihre Bestandteile zerlegt. Dabei zeigte sich, dass Fraktionen, welche nur das Protein oder die RNA enthielten, nicht enzymatisch aktiv waren. Allerdings konnte die Aktivität bei der gleichzeitigen Inkubation von RNA und Protein wiederhergestellt werden. Zusätzlich konnte gezeigt werden, dass die RNA-Komponente auch *in vivo* essenziell ist [4, 27]. Zu Beginn der 1980er-Jahre wurde erstmals gezeigt, dass die RNA-Untereinheit der RNase P, welcher zunächst nur Hilfsfunktionen zugeschrieben wurden, unter Hochsalzbedingungen die Abspaltung der 5'-Flanke auch ohne Proteinuntereinheit durchführen kann und folglich ein sogenanntes „Ribozym“ darstellt [28]. In den nächsten Jahren wurde die RNA-basierte RNase P immer weiter untersucht, sowie strukturell aufgelöst (Abbildung 2) [29-32]. Zusätzlich zu den *in vitro* Studien konnte auch gezeigt werden, dass die bakterielle RNase P in den verschiedenen Organismen durch weitverwandte P RNAs oder P Proteine ersetzt werden kann [33-36]. Neben den bakteriellen RNase P RNAs wurden auch in den anderen Domänen des Lebens, den Archaea und Eukarya, aktive P RNAs nachgewiesen [37, 38]. Im Gegensatz zu den meisten bis heute bekannten Ribozymen, die sich meistens nur selbst spalten können, ist die RNase P zu einem mehrfachen Substratumsatz *in trans* fähig, da diese bei der Hydrolyse nicht beschädigt wird [39, 40]. Für die Entdeckung katalytisch-aktiver

RNA wurden S. Altman und T. Cech 1989 mit dem Nobelpreis für Chemie ausgezeichnet.

Um die Jahrtausendwende wurde die Vermutung aufgestellt, dass es auch noch eine andere Form der RNase P geben könnte. So zeigten die RNase P aus den menschlichen Mitochondrien, sowie aus Spinatchloroplasten eine proteinähnliche Dichte und waren gegenüber einer spezifischen Methode zum Nachweis von Ribozymen resistent [41-44]. Diese Vermutung wurde als erstes bestätigt, indem die humane mitochondriale RNase P gereinigt und charakterisiert wurde. Es zeigte sich, dass diese RNase P frei von einer RNA Untereinheit ist und aus drei Proteinuntereinheiten (MRPP1-3) besteht [45]. Kurz darauf wurden Orthologe der funktionalen Untereinheit MRPP3 auch in Pflanzen nachgewiesen. Es zeigte sich, dass in Pflanzen lediglich das MRPP3 für die Spaltung der Prä-tRNA notwendig und somit als singuläre Protein-basierte RNase P (PRORP) existent ist [46, 47]. Dieses Protein ist auch in der Lage, die RNase P von *E. coli* *in vivo* zu ersetzen [46]. Neben dem Vorkommen in Pflanzen konnte diese Protein-basierte Form auch in dem einzelligen eukaryotischen Parasiten *Trypanosoma brucei* nachgewiesen werden. Dieses PRORP-Enzym ist in der Lage, die nukleäre RNase P von Hefe (eine RNA plus neun Proteine) *in vivo* funktional zu ersetzen [48]. Nachdem es schien, dass Protein-basierte RNase Ps exklusiv in Eukaryoten vorkommen, konnte 2017 gezeigt werden, dass das Bakterium *Aquifex aeolicus* lediglich ein kleines Protein zur Prozessierung der Prä-tRNAs einsetzt. Homologe der RNase P von *A. aeolicus* (HARPs) konnten ausschließlich in Bakterien und Archaeen nachgewiesen werden [49]. Diese Entdeckung erweitert das Feld der Protein-basierten RNase P auf alle drei Domänen des Lebens, wie es auch für die RNA-basierte RNase P gezeigt wurde. Zusätzlich zu einer Vielzahl funktionaler Studien konnten auch mehrere Strukturen für PRORP aufgeklärt werden (Abbildung 3) [50-52].

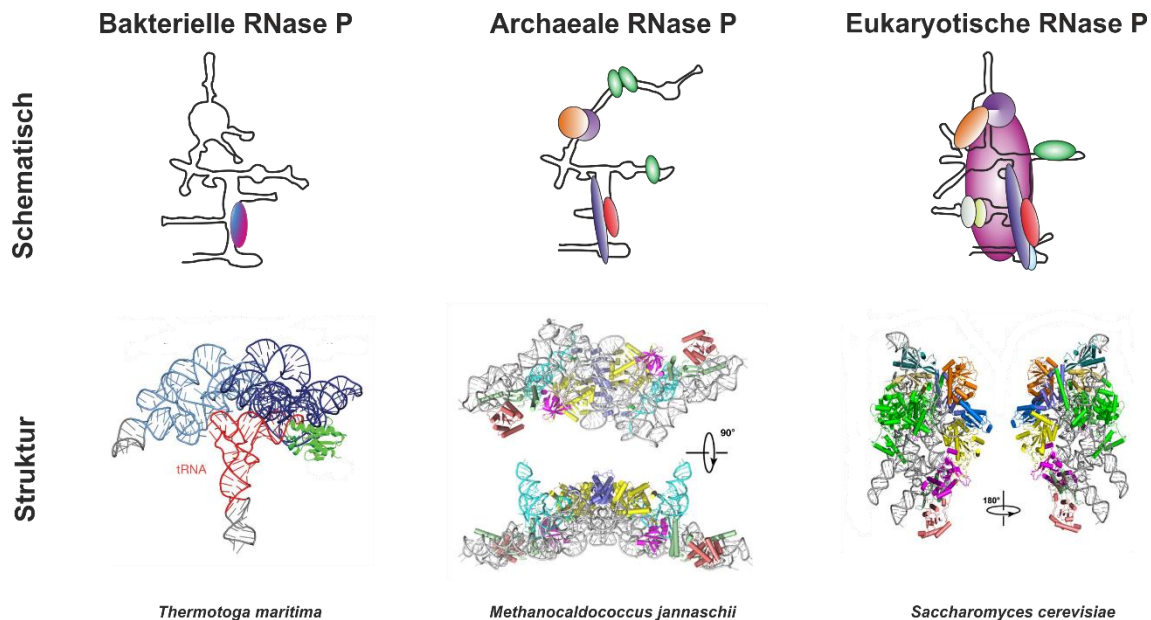
### 1.2.1 Die RNase P als Ribozym

Die RNase P besteht in Bakterien aus einer katalytischen RNA mit 350-400 Nukleotiden (nt), sowie einer kleinen Proteinuntereinheit (13-14 kDa). Die RNase P ist ein essenzielles Enzym und nicht kann deletiert werden, ohne in einem letalen Phänotyp zu enden. Allerdings gibt es einige wenige

temperatursensitive Mutanten, welche die tRNA-Prozessierung betreffen [25, 53]. Die RNA Untereinheit der RNase P (P RNA) ist strukturell und funktionell untersucht worden [54-57] und kann in basierend auf strukturellen Unterschieden in verschiedene Klassen unterteilt werden [57, 58]. Demnach wird zwischen dem Typ A (*ancestral type*) und Typ B (*Bacillus type*) differenziert [59]. Ferner kann die P RNA in zwei Domänen unterteilt werden (Abbildung 4C), welche sich unabhängig voneinander falten können. Die S-Domäne interagiert mit dem T-Arm der Prä-tRNA und ist damit in die Substratspezifität involviert. Die C-Domäne beinhaltet das aktive Zentrum, sowie die Interaktionsfläche für das P Protein und vermittelt somit die Katalyse des Enzyms. Trotz dieser Unterschiede existieren in Bakterien fünf konservierte Bereiche der P RNA (CR I-V) [59-61]. Es konnte gezeigt werden, dass verschiedene bakterielle P RNAs in der Lage sind *in vivo* eine funktionelle Komplementation hervorzurufen [33, 34]. *In vitro* kann die RNA Untereinheit unter Hochsalzbedingungen allein die Prozessierung der Prä-tRNA durchführen [28, 32].

Die Proteinuntereinheit (P Protein) der bakteriellen RNase P besitzt mehrere Aufgaben. Sie dient der Stabilisierung der Tertiärstruktur der P RNA, der Erhöhung der Substratspezifität, sowie hilft bei der Dimerisierung des Holoenzym [28, 62-69]. Die P Proteine der Bakterien sind auch zwischen entfernt verwandten Familien strukturell sehr ähnlich [70-72] und können untereinander ersetzt werden [73]. Die Funktionen der P Proteine der Archaea und Eukarya ist noch nicht ganz verstanden. Sie dienen sehr wahrscheinlich ebenfalls der Stabilisierung, indem sie fehlende RNA-Strukturen ersetzen, die bei bakterieller P RNA für Tertiärstruktur-Wechselwirkungen wichtig sind [74, 75]. In Abbildung 2 sind die verschiedenen RNase P Typen dargestellt. Dabei sind in der oberen Reihe die Strukturen nur schematisch abgebildet. Es ist zu erkennen, dass mit steigender Komplexität des Organismus (Bakterien ► Eukarya) die Zahl der Proteinuntereinheiten ebenfalls zunimmt. Bakterien haben nur eine, Archaea bis zu fünf und Eukarya bis zu zehn P Proteinuntereinheiten. Einige der P Proteinuntereinheiten in Eukarya und Archaea sind homolog zueinander. In der unteren Reihe sind die aufgelösten Strukturen zum Teil im Komplex mit einer tRNA dargestellt. Dabei ist die RNA in Strichweise, sowie die Proteine als Cartoon abgebildet. Es zeigte sich, dass das bakterielle P Protein mit der 5'-

Flanke der Prä-tRNA interagiert. In den komplexeren Organismen nimmt die Interaktion zwischen P Proteinen und tRNA zu (siehe 1.3).



**Abbildung 2: Schematische und strukturelle Darstellungen RNA-basierter RNase Ps**

In der oberen Reihe sind exemplarisch die verschiedenen Formen der RNA-basierten RNase P schematisch dargestellt. In Schwarz ist die P RNA und in Farbe die verschiedenen P Proteine dargestellt. Dabei spiegeln identische Farben homologe Proteine zwischen den Domänen wider. Das Protein der Bakterien weist keine Homologen auf. In der unteren Reihe sind aufgelöste Strukturen dargestellt. In allen Strukturen ist die RNA-Untereinheit als Strich-Modell und die Proteine als Cartoon-Modell gezeichnet. Die Farben der Proteine in der strukturellen Darstellung entsprechen nicht den Farben in der schematischen Darstellung. Stellvertretend für Bakterien handelt es sich dabei um die Struktur der RNase P aus von *Thermotoga maritima* im Komplex mit einer tRNA (rot) [71]. Für die Archaea ist die Cryo-EM Struktur einer dimeren RNase P aus *Methanocaldococcus jannaschii* [76], ebenfalls im Komplex mit zwei tRNAs (türkis) dargestellt. Für die Eukarya ist exemplarisch die Struktur der RNase P aus Bäckerhefe (*Saccharomyces cerevisiae*) [77] abgebildet.

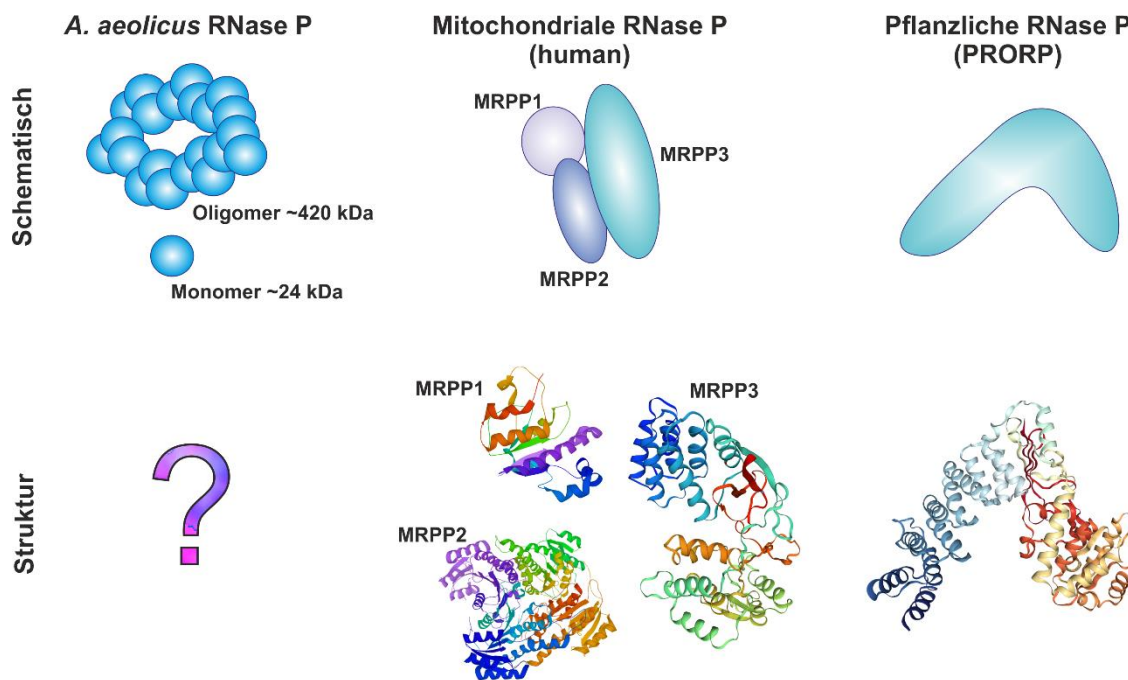
### 1.2.2 Die RNase P als (singuläres) Protein

Um die Jahrtausendwende wurde die Vermutung aufgestellt, dass neben der RNA-basierten RNase P noch eine weitere Form vorliegen könnte, welche ausschließlich aus Proteinen bestehen könnte [41-44]. Dies konnte durch die Charakterisierung der RNase P aus humanen Mitochondrien, welche aus drei Proteinuntereinheiten (MRPP1-3) besteht, bestätigt werden [42, 45]. Später konnte in der Landpflanze *Arabidopsis thaliana* ein RNase P-Enzym (PRORP)

identifiziert werden, bei dem es sich um ein singuläres Protein handelt [46] (Abbildung 3). Schließlich gelang auch in Bakterien die Charakterisierung eines Protein-basierten RNase P-Enzyms: in *Aquifex aeolicus* ist die RNase P ebenfalls ein singuläres Protein, bei dem es sich um eine minimalisierte Form von PRORP handelt [49].

Die humane mitochondriale RNase P besteht, aus drei Untereinheiten (MRPP1-3). Dabei bildet die tRNA m<sup>1</sup>G Methyltransferase MRPP1 einen stabilen Komplex mit der Hydroxysteroid(17- $\beta$ )Dehydrogenase MRPP2 [45, 78, 79]. Beide Untereinheiten zeigen keine RNase P Aktivität. Diese wird erst durch die Untereinheit MRPP3, welche das katalytische Zentrum enthält vermittelt [45]. Diese Untereinheit der humanen mitochondrialen RNase P hat einige Homologe in anderen Eukarya, unter anderem das singuläre Protein PRORP aus *A. thaliana* [42, 45-47, 80]. In *A. thaliana* existieren drei verschieden lokalisierte PRORP-Isoenzyme. PRORP1 liegt in den Organellen vor, wohingegen PRORP2 und PRORP3 beide im Kern detektiert werden können [47]. Während die Deletion von PRORP1 zu einem letalen Phänotyp führt, kann bei einer Deletion von PRORP2 oder PRORP3 kein Unterschied zum Phänotyp des Wildtyps festgestellt werden. Demnach haben die beiden nukleär lokalisierten Enzyme anscheinend eine redundante Funktion [46, 47]. Auch in anderen höheren Organismen konnten mehrere PRORPs nachgewiesen werden, so hat das Laubmoos *Physcomitrella patens* ebenfalls drei verschieden lokalisierte Isoenzyme, wobei nur ein Enzym im Kern lokalisiert ist und die andern zwei in den Organellen [81].

Alle eukaryotischen PRORPs, sowie auch MRPP3, gehören zu der Familie der Metallonukleasen, welche durch eine „NYN“-Domäne charakterisiert werden [82] und ein *pentatricopeptide repeat* (PPR)-Motiv aufweisen [83]. Das PPR-Motiv besteht aus 35 Aminosäuren, welche bis zu 30-mal wiederholt werden können und unter anderem in der RNA Bindung eine Rolle spielen [83, 84]. Die Struktur von PRORP1 ist aufgelöst weist eine „V“-Form auf. Dabei gliedert sich das Protein in drei Bereiche, den bereits genannten PPR-Motiv (blau) und der NYN-Metallonuklease-Domäne mit dem katalytischen Zentrum (rötlich). Beide Domänen werden durch das zentrale Zink-Finger Motiv (eisblau) getrennt, welches die „V“-Form vermittelt indem es die Bindungs- und Katalyse-Domäne zueinander koordiniert [51] (Abbildung 3).



**Abbildung 3: Schematische und strukturelle Darstellungen Protein-basierter RNase P**

In der oberen Reihe sind exemplarisch die verschiedenen Formen Protein-basierter RNase P schematisch dargestellt. Bei der Darstellung der RNase P von *A. aeolicus* handelt es sich um eine Interpretation experimentell gewonnener Daten wobei die Anordnung der Monomere im Oligomer rein hypothetisch ist. Eine aufgelöste Struktur ist nicht existent. Die humane mitochondriale RNase P ist aus drei Untereinheiten aufgebaut (MRPP1-3), wobei MRPP3 die Spaltungsreaktion durchführt. Die Strukturen der Proteinuntereinheiten wurden einzeln aufgelöst [52, 85, 86], eine Komplexstruktur ist nicht verfügbar. Bei der pflanzlichen RNase P (PRORP) handelt es sich um ein Enzym von *A. thaliana*. Es ist homolog zur Untereinheit MRPP3 der humanen mitochondrialen RNase P. Die aufgelöste Struktur ist von At\_PRORP1 [51].

### 1.3 Substraterkennung und Prozessierung durch RNase P

Die Substraterkennung unterscheidet sich zwischen RNA- und Protein-basierten RNase P Enzymen. Aufgrund der unterschiedlichen Strukturen, Größe und dem chemischen Aufbau des Enzyms sind unterschiedliche Bereiche der Prä-tRNA relevant für die Erkennung. Ein detaillierterer Einblick in die Substraterkennung wird in den nachfolgenden Kapiteln aufgezeigt.

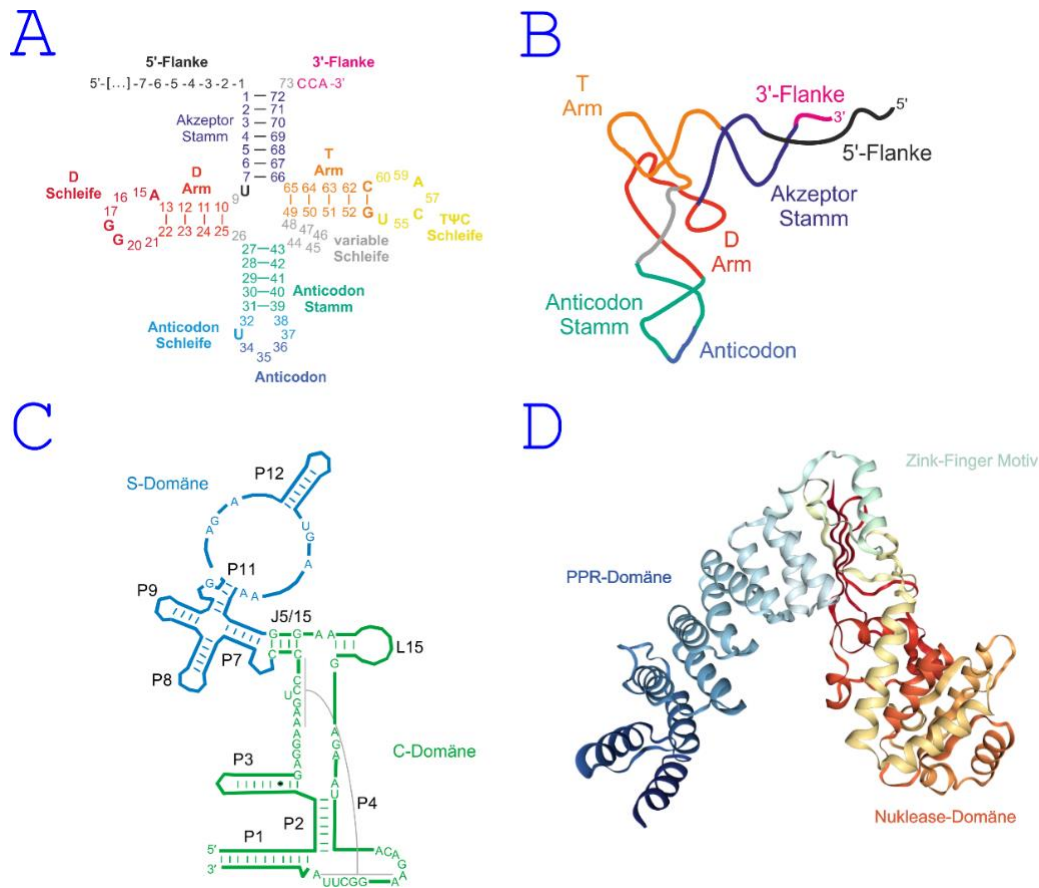
Im Gegensatz zur Substratbindung ist die durch das Ribozym und das Protein-basierte Enzym katalysierte Reaktion ähnlich. In beiden Fällen handelt es sich um eine endonukleolytische Hydrolyse einer Phosphodiesterbindung. Dabei entsteht ein 5'-Phosphat an der prozessierten tRNA und eine 5'-Flanke mit einer 3'-OH Gruppe [59, 87]. Der Reaktionsmechanismus ist stark abhängig von zweiwertigen Kationen, vornehmlich Magnesium ( $Mg^{2+}$ ) [61, 88].



### 1.3.1 RNA-basierte RNase P

Die Erkennung der verschiedenen Substrate muss über einen allgemeinen Mechanismus erfolgen, bei dem weniger die Sequenz als die Struktur eine Rolle spielt. Die Sequenzen der tRNAs sind bis auf wenige Ausnahmen (Abbildung 4A) nicht konserviert, wohingegen fast alle kanonischen tRNAs eine gemeinsame Tertiärstruktur aufweisen (Abbildung 4B). So sind das 5'- und 3'-Ende durch einen 7 bp langen Akzeptorstamm miteinander verbunden, wobei am 5'-Ende eine Vorläuferflanke und am 3'-Ende das CCA-Motiv übersteht (Abbildung 4A und B). Zwei Ausnahmen sind die tRNA<sup>His</sup> und tRNA<sup>SeC</sup> [89], welche eine Verlängerung des Akzeptorstamms aufweisen. Im bakteriellen System ist die Länge des Akzeptorstamms relevant, spielt eine Rolle bei der Erkennung der kanonischen Spaltstelle +1/-1 und interagiert mit dem U<sub>69</sub> der P RNA [90-92]. Die Vorläuferflanke (5'-Flanke) interagiert im Gegensatz zum tRNA Körper nicht mit der P RNA, sondern nur mit dem P Protein. Dabei sind die Nukleotide -2, sowie -4 bis -7 involviert [66, 93, 94]. Zwar ist auch die P RNA allein in der Lage eine Prä-tRNA zu prozessieren, jedoch wird die Affinität zum Substrat durch die Proteinuntereinheit signifikant erhöht [66, 95, 96]. Ein weiterer Erkennungsmechanismus ist die Interaktion zwischen der TΨC-Schleife der Prä-tRNA mit den P7-P11 Schleifen der P RNA [97-99]. Eine untergeordnete Rolle in der Substraterkennung spielen der Anticodon-Stamm, sowie der D-Arm mit der Dihydrouracilschleife (D-Schleife) [100]. Eine weitere Interaktion findet zwischen dem 3'-CCA Ende der tRNA statt, welches in vielen Organismen codiert auf dem Genom vorliegt und den Nukleotiden in der C-Domäne der P RNA statt. Konkret kommt es zu Watson-Crick-Basenpaarungen zwischen den konservierten Nukleotiden GG(U) in der Schleife J15/16 (L15 für Typ B) der P RNA und den komplementären 3'-(R)CCA-Ende ausgebildet [101]. Unter allen diversen Wechselwirkungen zwischen dem Substrat und der RNase P stellen die Interaktionen, bei der Nukleotide in der Helix und des Akzeptorstamms und des T/D-Arm der tRNA involviert sind, die primären Interaktionen dar [91, 102, 103]. Im archaealen System werden die Hauptinteraktionen ebenfalls durch die P RNA vermittelt. Dabei sind die Wechselwirkungen zwischen der D-Schleife, der TΨC-Schleife und den konservierten Regionen II und III der P RNA, sowie die Interaktion mit dem Akzeptorstamm ähnlich zu dem bakteriellen System [76,

104]. Die Interaktion zwischen dem 3'-(R)CCA Ende mit einem konservierten Bereich der RNA Untereinheit scheint nicht vorhanden zu sein [76]. Die Proteinuntereinheiten Rpp21-Rpp29 haben zudem einen Einfluss auf die Bindung, sowie die Spaltstellenbestimmung [104].



**Abbildung 4: Übersicht von Prä-tRNA-Strukturen und RNase P-Domänen**

In dieser Abbildung sind die Primär- und die Tertiärstruktur einer Prä-tRNA abgebildet. Zusätzlich zu den Substraten ist Sekundärstruktur einer bakteriellen Konsensus- P RNA, sowie die Tertiärstruktur von PRORP1 dargestellt. (A) Sekundärstruktur einer Prä-tRNA mit farblichen Kennzeichnungen der einzelnen Bereiche. Die konservierten Nukleotide sind ausgeschrieben (A, G, C oder U). Die Prä-tRNA setzt sich zusammen aus der 5'-Flanke, dem Akzeptorstamm (i.d.R. 7 nt), dem D-Arm mit der Dihydrouracilschleife (D-Schleife), dem T-Arm (5 nt) mit der TΨC-Schleife (7 nt), sowie dem Anticodon-Stamm (7 nt) mit der Anticodon-Schleife und dem Anticodon in der Mitte, dem 3'-RCCA-Ende, sowie einer variablen Schleife (nach [89, 100, 105]). (B) Tertiärstruktur einer Prä-tRNA, in welcher die typische L-Form deutlich wird. Die farbliche Kennzeichnung der einzelnen Bereiche entspricht Abbildung (A). Es ist zu erkennen, dass der T-Arm den „Ellbogenbereich“ bildet (nach [89, 100, 105]). (C) Minimale Konsensus Struktur einer bakteriellen RNase P RNA. Hervorgehoben sind die einzelnen Schleifen, sowie farblich die Spezifitäts- (S-Domäne) und die katalytische Domäne (C-Domäne). Konservierte Nukleotide sind ausgeschrieben (A, C, G oder U) (angelehnt an [106]). (D) Kristallstruktur der Protein-basierten RNase P von *A. thaliana* (PRORP 1). Es besteht aus einer PPR-Domäne, (pentatricopeptide repeat), welche an der RNA Bindung beteiligt ist, einem zentralen Zink-Finger Motiv, sowie einer Nuklease Domäne, welche in der Prozessierung der Substrate und Koordination der Metallionen eine Rolle spielt [51].



### 1.3.2 Protein-basierte RNase P

Die Substraterkennung der Protein-basierten RNase P Enzyme (PRORP) ist am besten an den pflanzlichen Vertretern aus *A. thaliana* untersucht worden. In *A. thaliana* sind drei unterschiedlich lokalisierte PRORP-Isoenzyme bekannt. PRORP1 ist in den Organellen (Mitochondrien und Chloroplasten) lokalisiert, wohingegen PRORP2 und PRORP3 sich im Nukleus befinden [47]. Es zeigten sich grundlegende Unterschiede in der Substraterkennung zwischen den Ribonukleoprotein-Enzymen und den PRORPs.

Weder die 5'-Flanke noch das 3'-Ende haben einen entscheidenden Einfluss auf die Substraterkennung und -prozessierung durch PRORP [107, 108]. Allerdings bevorzugt PRORP2 kurze flankierende Sequenzen an beiden Enden [50]. Die Substratbindung von maturer tRNA war signifikant schlechter als die von Prä-tRNA mit 5'-Flanke [107]. Somit ist der tRNA Körper für die Erkennung und Bindung relevant. PRORP1 und PRORP2 zeigten mindestens vier direkte Interaktionen mit dem tRNA Körper [109]. Es zeigte sich bei *Footprinting* Experimenten mit PRORP1, dass Bereiche der D- und TΨC-Schleife geschützt wurden [110]. In späteren Versuchen wurde deutlich, dass die TΨC-Schleife, zusammen mit dem Akzeptorstamm für eine Substratbindung ausreichend ist, aber die Affinität mit dem D-Arm erhöht werden kann [107, 108]. Der Anticodon-Stamm scheint bei der Substraterkennung nicht relevant zu sein [108]. Ein *hairpin*-Minimalsubstrat, welches einem koaxial gestapelten Akzeptor-/T-stamm mit TΨC-Schleife entspricht, wird allerdings bedeutend schlechter gebunden und prozessiert [107, 108]. Wurden in diese Stammstruktur einzelne Nukleotide in die 5'-Hälfte der Helix eingeführt (*bulge* von 9 Nukleotiden) eingeführt, so dass sich eine leicht L-förmige Struktur ausbilden konnte, wurde die Prozessierung deutlich verbessert [108]. In den Enzymen scheint für die Spaltgenauigkeit die Distanz zwischen Akzeptorstamm und dem T-Arm mit Schleife relevant zu sein, denn eine Verlängerung, sowie Verkürzung führt zu einer Fehlsplaltung [108, 110, 111].

## 1.4 *Aquifex aeolicus*

*Aquifex aeolicus* ist ein hyperthermophiles, motiles, nicht-sporulierendes und Gram-negatives Bakterium, welches zur Familie der *Aquificaceae* gehört. Mit

einer Wachstumstemperatur von bis zu 95°C zählt es zu den thermophilsten Organismen, die bisher bekannt sind (mehr Informationen unter Publikation III im Anhang).

Diese Familie besitzt keine Gene für die Untereinheiten der bakteriellen RNase P im Gegensatz zu den Schwesterfamilien. Allerdings besitzen diese Bakterien Gene für eine neuartige Form der RNase P. Diese ist Protein-basiert und nur ungefähr 23 kDa groß [49]. Entsprechend ihrer Herkunft wurde diese neue RNase P-Form AaRP (*Aquifex aeolicus* RNase P) benannt.

## 2 Zielsetzung

Die RNase P ist ein essenzielles Enzym, welches in allen drei Domänen des Lebens identifiziert werden konnte. Es ist verantwortlich für die Spaltung der 5'-Flanke von Vorläufer-tRNAs. Es sind zwei Formen dieses Enzyms bekannt, die RNA-basierte Form, welche aus einer katalytischen RNA und ein bis zehn Proteinuntereinheiten besteht, sowie einer rein Protein-basierten Form, welche zunächst nur in höheren Organismen identifiziert werden konnte.

In dieser Arbeit sollte sowohl die Identifizierung der RNase P aus *Aquifex aeolicus* als auch die Charakterisierung erfolgen. Zunächst wurden die in Vorarbeiten identifizierten P Protein-Kandidaten rekombinant hergestellt und biochemisch untersucht. Es stellte sich heraus, dass die RNase P von *A. aeolicus* (AaRP) als singuläres Protein vorliegt. Durch dieses Ergebnis ergaben sich drei weitere Forschungsprojekte, welche sich alle mit der prokaryotischen Protein-basierten RNase P beschäftigten. Zunächst sollte die neu identifizierte Form der RNase P genauer charakterisiert werden. Der Fokus sollte dabei auf der Substratspezifität, der Substratbindung, der enzymatischen Aktivität sowie einer vermeintlichen Oligomerisierung liegen. Dies sollte sowohl durch *in vitro* Studien mit dem rekombinanten Protein als auch *in vivo* in verschiedenen bakteriellen Modellorganismen untersucht werden. Das in *A. aeolicus* identifizierte Enzym ist die bisher kleinste bekannte Form einer Protein-basierten RNase P und weist erstaunlicherweise keine erkennbare Substratbindungs-Domäne auf. Durch Kristallisationsexperimente sollte die Struktur dieses neuartigen RNase P-Enzyms aufgeklärt werden, was eventuell auch Erkenntnisse dazu liefern könnte, wie das Enzym Substratspezifität erlangt.

Im Verlauf dieser Studien wurden durch bioinformatische Analysen Homologe von AaRP (HARP) in verschiedenen Prokaryoten identifiziert. Erstaunlicherweise wiesen diese Organismen sowohl Gene für die RNA-basierte RNase P als auch für das neu entdeckte HARP auf. Daraus entstanden zwei Nebenprojekte, die sich mit der Fragestellung beschäftigten, welche Funktion die HARP-Enzyme in den jeweiligen Organismen besitzen und ob sie tatsächlich auch RNase P-Aktivität aufweisen. Weiterhin sollten von den Archaeen und Bakterien stellvertretend einige Organismen genauer bezüglich ihrer RNase P-Enzyme analysiert werden, um mögliche Unterschiede zwischen archaealen und bakteriellen HARP-Varianten aufzudecken. Zusätzlich zu den Studien mit den rekombinanten Enzymen sollten die archaealen HARPs auch im jeweiligen ursprünglichen Organismus (*Haloferax volcanii* und *Methanosarcina mazei*) funktionell analysiert werden. Bei den ausgewählten Bakterien mit beiden RNase P-Formen wurden sowohl die HARPs als auch die klassischen RNA-basierten RNase P-Enzyme analysiert. Da bei den zum größten Teil extremophilen Bakterien Studien im ursprünglichen Organismus nicht möglich waren, sollten sowohl die bakteriellen HARPs als auch die beiden Untereinheiten der RNA-basierten RNase P (kodiert von *rnpA* und *rnpB*) für *in vivo*-Studien in *E. coli* und *B. subtilis* exprimiert werden.

### 3 Material und Methoden

#### 3.1 Bakterienstämme

Die folgende Tabelle enthält alle in dieser Arbeit verwendeten Bakterienstämme. Zusätzlich zu den Gattungsnamen sind auch der Genotyp, die Referenz sowie die entsprechende Inkubationstemperatur angegeben. Für die Vermehrung klonierter Vektoren wurde das Gram-negative Bakterium *Escherichia coli* (*E. coli*) DH5 $\alpha$  verwendet. Für die Expression rekombinanter Proteine wurden *E. coli* BL21(DE3) oder *E. coli* Rosetta(DE3) genutzt. Die Komplementation der endogenen RNase P durch eine Vektor-exprimierte RNase P wurde in *E. coli* BW durchgeführt. Der Gram-positive Stamm *Bacillus subtilis* (*B. subtilis*) d7 wurde zur Analyse der Funktionalität des P Proteins der RNase P *in vivo* verwendet.

*Tabelle 1: Auflistung der in dieser Arbeit verwendeten Bakterienstämme.*

Es ist der Organismus, sowie der Genotyp, die Inkubationstemperatur und die Referenz angegeben.

<i>E. coli</i>	Genotyp	Temp.	Referenz
DH5 $\alpha$	F <sup>-</sup> endA1glnV44thi-1recA1relA1gyrA96deoRnupGpurB20 $\phi$ 80dlacZ $\Delta$ M15 $\Delta$ (lacZYA-argF)U169, hsdR17( <i>r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup></i> ), $\lambda$ <sup>-</sup>	37°C	[112]
BL21(DE3)	F <sup>-</sup> ompTgaldcmlonhsdS <sub>B</sub> ( <i>r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup></i> ) $\lambda$ (DE3 [lacI lacUV5-T7p07ind1sam7nin5]) [ <i>malB</i> <sup>+</sup> ] <sub>K-12</sub> ( $\lambda$ <sup>S</sup> )	37°C	NEB
Rosetta(DE3)	F <sup>-</sup> ompThsdS <sub>B</sub> ( <i>r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup></i> ) gal dcm (DE3) pLysSRARE (Cam <sup>R</sup> )	30°C-37°C	Novagen
BW		26°C-37°C	[113]
BL21(DE3)pLy sS	F <sup>-</sup> ompTgaldcmlonhsdS <sub>B</sub> ( <i>r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup></i> ) $\lambda$ (DE3[lacI lacUV5-T7p07ind1sam7nin5]) [ <i>malB</i> <sup>+</sup> ] <sub>K-12</sub> ( $\lambda$ <sup>S</sup> )	37°C	NEB
<i>B. subtilis</i>	Genotyp	Temp.	Referenz
d7	trpC2 metB10 SP $\beta$ - (Eyr), rnpA::P <sub>xylose</sub>	28-37°C	[114]

### 3.1.1 Lagerung und Kultivierung der Bakterienstämme

Die in Tabelle 1 gelisteten Stämme wurden für eine Langzeitlagerung bei -80°C als 50%-ige Glycerol-Kultur angelegt. Für eine Lagerung bis maximal 4 Wochen wurden der gewünschte Stamm auf eine 1%-ige LB-Agar Platte ( $\pm$  Antibiotika;  $\pm$  Kohlenstoffquelle) ausgestrichen, bis zu einer Koloniebildung bei entsprechender Temperatur inkubiert und anschließend entweder bei Raumtemperatur (*B. subtilis*) oder 4°C (*E. coli*) gelagert.

Für die Vermehrung eines Klons in Flüssigmedium wurden für analytische Untersuchungen 3 - 5 mL oder für Expressionen 250 - 1000 mL Flüssigkultur mit den entsprechenden Antibiotika beimpft und bei 180 - 220 rpm (GSL-3300, Luftinkubator) inkubiert. Es wurden folgende Antibiotika einzeln oder in verschiedenen Kombinationen verwendet: 100  $\mu$ g/mL Ampicillin, 50  $\mu$ g/mL Kanamycin oder 34  $\mu$ g/mL Chloramphenicol für *E. coli*, sowie 20  $\mu$ g/mL Kanamycin und 5  $\mu$ g/mL Chloramphenicol für *B. subtilis* d7. Bei Verwendung von *E. coli* BW wurde zusätzlich 10 mM Arabinose als Kohlenstoffquelle zugegeben. Bei Zugabe von 10 mM Glukose wurden Komplementationsexperimente (3.7.1) mit *E. coli* BW durchgeführt. Im Falle von dem Stamm *B. subtilis* d7 wurde dem Medium entweder 2% Xylose (v/v) oder 2% Glukose (v/v) zugegeben und ebenfalls Komplementationsexperimente durchgeführt.

### 3.1.2 Herstellung chemisch-kompetenter Bakterien

Chemisch-Kompetente Bakterien werden benötigt, um genetisches Material durch eine Thermotransformation aufnehmen zu können. In dieser Arbeit wurden verschiedene *E. coli*-Stämme durch zwei unterschiedliche Methoden Kompetent gemacht.

Im Falle von *E. coli* DH5 $\alpha$  wurde eine 3 mL Übernachtskulturen beimpft und damit am nächsten Tag eine Kultur mit einer Optischen Dichte (OD<sub>600</sub>) von 0.1 beimpft. Die Kultur wurde bis zu einer OD<sub>600</sub> von 0.5 - 0.6 inkubiert. Nach dem Erreichen wurden die Zellen für 10 - 15min auf Eis inkubiert und anschließend geerntet (7 min, 5 000 rpm, 4°C). Je 50 mL Kultur wurde das Pellet in 15 mL eiskalter TFB-1 Lösung aufgenommen und anschließend wieder geerntet. Pro 50 mL Ausgangskultur wurde das Pellet nun in 6 mL TFB-3 Lösung

aufgenommen und in eisgekühlte Reaktionsgefäße aufgeteilt. Die aliquotierten Zellen wurden in flüssigen Stickstoff eingefroren und bis zur Verwendung bei -80°C gelagert.

Für *E. coli* BW, Rosetta (DE3), BL21 (DE3) und BL21 (DE3) pLysS wurde eine Übernachtskultur am nächsten Morgen 1:100 in frisches LB-Medium verdünnt und bis zu einer OD<sub>600</sub> von 0.2-0.5 weiter inkubiert. Anschließend wurden die Zellen für 10 min auf Eis gelagert, 3 min bei 5 000 rpm und 4°C geerntet und das Zellpellet in 1/10 V des Ursprungsvolumens in TSS-Puffer (5 g PEG 8000, 1.5 mL MgCl<sub>2</sub> (1 M), 2.5 mL DMSO auf 50 mL mit LB) aufgenommen. Diese Zellen wurden entweder sofort für eine Transformation genutzt oder in 50 µl Aliquot bei -80°C gelagert.

### 3.1.3 Herstellung elektro-kompetenter Bakterien

Für die Elektrottransformation wurden *B. subtilis* Zellen verwendet, welche zuvor kompetent gemacht wurden. Dazu wurde ein Protokoll verwendet, welches durch die verwendeten Puffer und Medien besonders die Stabilisierung der transformierten Zellen begünstigt (angelehnt an [115]).

### 3.1.4 Transformation von genetischem Material (Hitzeschock)

Das für die Thermotransformation genutzte genetische Material stammte entweder aus einer Ligation (3.2.5) oder es wurde ein bereits klonierter und sequenzierter Vektor (Tabelle 4) genutzt. Das folgende Protokoll (angelehnt an [116]) wurde sowohl für die Transformation von *E. coli* DH5α, als auch *E. coli* BW und *E. coli* Rosetta (DE3) benutzt. Dabei variierte lediglich der erste Schritt des Protokolls, je nachdem um welches genetische Ausgangsmaterial es sich handelte.

Im Falle einer Ligation (3.2.5) wurden 20 µl des Ansatzes mit 30 µl chemisch-kompetenten *E. coli* DH5α gemischt und für 15 - 30 min auf Eis inkubiert. Handelte es sich um einen bereits klonierten und sequenzierten Vektor wurden 200 - 500 ng bei einem maximalen Volumen von 5 µl mit 50 µl chemisch-kompetenten Zellen gemischt und für 5 - 10 min auf Eis inkubiert. Anschließend wurde die Zellsuspension für 1,5 min auf 42°C erhitzt, kurz auf Eis abgekühlt und

mit 500 µl LB-Medium ohne den Zusatz von Antibiotika versetzt. Für die Regeneration wurden die Zellen für 45 - 60 min, schüttelnd bei 37°C inkubiert bevor sie geerntet (2 min, 8 000 rpm) wurden. Das Pellet wurde in maximal 70 µl LB Medium aufgenommen und danach auf die entsprechenden Selektionsplatten ausplattiert und über Nacht bei 37°C inkubiert.

### 3.1.5 Transformation von genetischem Material (Elektroporation)

Die Elektroporation wurde verwendet, um *B. subtilis* zu transformieren. Dazu wurde ein bereits klonierter und sequenzierter Vektor (Tabelle 4) verwendet, vermehrt in *E. coli* DH5α und gereinigt.

Die elektrokompenten Zellen (80 µL) wurden mit mindestens 500 ng Plasmid-DNA gemischt, für einige Minuten auf Eis inkubiert und in eine Elektroküvette (0,2 cm) überführt. Die Transformation erfolgte mit Hilfe des Pulsgenerators (BioRad).

Nach dem Einzelimpuls (2,5 kV, 50 µF, 100 Ω) wurden die Zellen mit 1 mL vorgewärmten Medium (37°C, *outgrowth*-Medium (LB, 0,5 M Sorbitol, 0,38 M Mannitol, 1% (w/v) Xylose)) vermischt, in ein Reaktionsgefäß überführt und für 2 h bei 37°C inkubiert. Die Zellen wurden geerntet (6 000 rpm, RT) in 80 µL restliches *outgrowth*-Medium aufgenommen und auf Selektionsplatten (LB-Medium, 2% (w/v) Xylose, kann 20, cam 5) ausplattiert. Die Koloniebildung erfolgte für 24 - 48 h bei 37°C.

### 3.1.6 Kolonie Polymerase-Ketten-Reaktion (PCR)

Die Kolonie-PCR wurde eingesetzt, um zum einen nach einer Ligation die Größe des gewünschten Inserts zu überprüfen, sowie nach einer Transformation, dass vorhanden sein des Vektors.

Für die Reaktion wurde der 5xHOTCytoPol-Mix (Cytogen, Deutschland) verwendet. Dabei wurde nach Herstellerangaben ein Reaktionsvolumen von 10 µl angesetzt und die Einzelkolonie, nach dem Transfer auf eine Masterplatte darin aufgenommen. Als Oligonukleotide wurden Sequenzierungsprimer, welche in Tabelle 2 aufgelistet sind verwendet. Das Thermoprotokoll befindet sich in Tabelle 3. Durch die Verwendung der Sequenzierungsprimer kann anhand der

Größe des PCR-Produktes auf das Vorhandensein des spezifischen DNA-Fragments geschlossen werden. Ohne ein Insert hat das Fragment des der Vektor pDG148(+) eine Größe von 440 bp und pET28a(+) von 317 bp.

Zur Analyse der Kolonie-PCR wurden 3.5 µl des Ansatzes durch ein 0.8%-iges Agarose-Gel (3.2.2) der Größe nach aufgetrennt. Wenn die Größe der Erwartung entsprochen hat, wurden diese Klone als analytische Kultur angezogen, entweder um eine Plasmidisolation durchzuführen oder um eine Glycerol-Kultur anzulegen.

*Tabelle 2: Auflistung der Oligonukleotide für die Kolonie-PCR*

In der nachfolgenden Tabelle sind die verwendeten Oligonukleotide dem entsprechenden Vektor zugeordnet. Die Sequenz ist in 5'-3'-Richtung angegeben und in Triplets aufgeteilt.

Vektor	Name	Sequenz 5'- 3'
<b>pDG148(+)</b>	pDG_Seq_for	GCT CAG CTG GAA TTC TAC AC
	pDG_Seq_rev	CTC AAA CAT CAA ATC TTA CAA ATG TAG TC
<b>pET28a(+)</b>	T7	TAA TAC GAC TCA CTA TAG GG
	T7 term	CTA GTT ATT GCT CAG CGG

*Tabelle 3: Thermoprotokoll für eine Kolonie-PCR*

In der nachfolgenden Tabelle ist das Thermoprofil für eine Kolonie PCR aufgeführt, welches für alle Konstrukte angewendet wurde.

95°C	15 min	} 25x
95°C	0.5 min	
<b>60°C</b>	0.5 min	
72°C	1.16 min	
72°C	5 min	
12°C	Pause	

### 3.2 Klonierung von Vektoren zur Charakterisierung rekombinanter Proteine

Um verschiedene Proteine gezielt zu exprimieren wurden die entsprechenden Gene in diverse Vektoren eingebracht, welche sich je nach Ihrer Nutzung in ihren



Eigenschaften unterschieden haben. Dabei wurden verschiedene Grundvektorarten verwendet: pJET1.2blunt (Thermo Scientific) oder pCR2.1TOPO (Invitrogen) für Zwischen-Klonierungen für den eigentlichen Zielvektor, pET28a(+) (Novagen) für die Expression der Proteine mit oder ohne einer Histidin-Fusion am N- oder C-terminalen Ende und pDG148(+) für die Analyse der Proteine in ihrer Fähigkeit der Komplementation von *E. coli* BW. Für einige Konstrukte wurden die Gene durch die Firma Biomatik (USA) synthetisiert und in pUC57 eingebracht. Diese wurden als Ausgangsmaterial für die Klonierungen in den entsprechenden Zielvektor verwendet. Eine Übersicht aller Kombinationen der zu untersuchenden Gene in den verschiedenen Vektoren finden sie in der nachfolgenden Tabelle 4.

### 3.2.1 Polymerase-Ketten-Reaktion (PCR)

Die Polymerase-Ketten-Reaktion (PCR) dient zur Vervielfältigung von spezifischen DNA-Sequenzen. Dabei wird folgender Zyklus mehrmals wiederholt: Im ersten Schritt wird der DNA-Doppelstrang denaturiert, als nächstes folgt die Anlagerung sequenzspezifischer Oligonukleotide und zum Schluss die Elongation durch eine thermostabile DNA-Polymerase.

Für die Amplifikation der Gene *aq\_880*, *aq\_707* und *aq\_221* aus genomischer DNA von *A. aeolicus* wurde der LongPCREnzyme Mix (Thermo Scientific) nach Herstellerangaben verwendet. Die Anlagerungstemperatur der Oligonukleotide lag bei 60°C [49]. Für die in Tabelle 4 aufgeführten Gene wurde sowohl der Phusion High-Fidelity PCR Master Mix with HF Buffer (Thermo Scientific), als auch die Invitrogen™Platinum™SuperFi DNA Polymerase (Invitrogen) nach Herstellerangaben verwendet. Die Anlagerungstemperatur der Oligonukleotide lag ebenfalls bei 60°C. Nach der Durchführung der PCR wurde der Erfolg und die Reinheit mittels einer Agarose-Gelelektrophorese (3.2.2) überprüft.

Tabelle 4: Übersicht der verwendeten Vektoren.

Auflistung der Vektoren zur Herstellung rekombinanter Proteine. Dabei wird sowohl die Genbezeichnung (Gen), als auch der Ursprungsorganismus (Organismus) angegeben. Die Gliederung der Vektoren pET28a(+) und pDG148(+) erfolgt in ein Fusionsprotein (His<sub>6</sub>) und ein nicht fusioniertes Protein (ohne). Die Vektoren des pGEX-mod besitzen am n-terminalen Ende einen GST (Glutathione S-transferase) Fusion, sowie eine Protease-Sequenz (PreScission, cut). Ein Häkchen (✓) spiegelt ein erfolgreich klonierten, ein Kreuz (✗) einen nicht klonierten Vektor wieder.

Gen	Organismus	pET28a(+)		pDG148(+)		pGEX-mod
		c-His <sub>6</sub>	ohne	c-His <sub>6</sub>	ohne	n-GST-cut
aq_880	<i>A. aeolicus</i>	✓	✓	✓	✓	✓
aq_221		✓	✗	✗	✗	✗
aq_707		✓	✗	✗	✗	✗
mth_243	<i>M. thermautotrophicus</i>	✓	✗	✓	✓	✗
thein_2060	<i>T. indicus</i>	✓	✗	✓	✓	✗
mtbma_c06930	<i>M. marburgensis</i>	✓	✗	✓	✓	✗
hvo_1114	<i>H. volcanii</i>	✓	✗	✓	✓	✗
minf_0007	<i>M. infernorum</i>	✓	✗	✓	✓	✗
hhal_2243	<i>H. halophila</i>	✓*n-His <sub>6</sub>	✗	✓	✓	✗

mlg_2556	<i>A. ehrlichii</i>	✓	✗	✓	✓	✗
tvnir_2893	<i>T. nitratireducens</i>	✓	✗	✓	✓	✗
mm_2077	<i>M. mazei</i>	✓	✗	✓	✓	✓
pfc_09185	<i>P. furiosus</i>	✓	✗	✓	✓	✓
thein_0634	<i>T. indicus</i>	✗	✗	✗	✓	✗
minf_0924	<i>M. infernorum</i>	✗	✗	✗	✓	✗
hhal_1229	<i>H. halophila</i>	✗	✗	✗	✓	✗
mlg_RS1457	<i>A. ehrlichii</i>	✗	✗	✗	✓	✗
tvnir_3809	<i>T. nitratireducens</i>	✗	✗	✗	✓	✗

### 3.2.2 Agarose-Gelelektrophorese

Die Agarose Gelelektrophorese ist eine Methode DNA-Fragmente nach der Größe auf zu trennen. Dabei spielt die Nukleotidsequenz gleich langer Fragmente keine Rolle, da das Masse-Ladungsverhältnis identisch bleibt.

Für die Agarose-Gelelektrophorese wurde ein 0.5 - 2% Agarosegel benötigt. Je niedriger die Agarose Konzentration des Gels, desto größere Fragmente wurden aufgetrennt. Für das Gel wurde die entsprechende Menge an Agarose eingewogen, mit 1x TBE aufgefüllt und aufgeköcht. Anschließend wurde dem aufgeköchten Gel GelRed (1:20000, Biotium) zugesetzt.

Standardmäßig wurden 5 µL einer DNA-Leiter aufgetragen. Die Gele wurden bei einer maximalen Spannung von 9-12 V/cm laufen gelassen. Nach dem Lauf wurden die Gele in einer Geldokumentationsanlage mittels ultravioletten Lichts dokumentiert.

### 3.2.3 Restriktionsverdau

Der Restriktionsverdau wird durch spezifische Nukleasen vermittelt. Diese erkennen in der Regel eine Sequenz von sechs Nukleotiden und spalten diese an unterschiedlichen Stellen. Dabei werden spezielle 5'- und 3'-Überhänge generiert. Wird sowohl der Vektor als auch das gewünschte Insert mit Nukleasen behandelt, die einen identischen 5'- und 3'-Überhang produzieren, kann man diese durch eine Ligation (3.2.5) mit einander verbinden.

Bei der Restriktion eines Vektors wurde 1 µg eingesetzt, bei einer PCR 10 µL des PCR-Ansatzes. Die Restriktionsenzyme wurden von Thermo Scientific bezogen und nach Herstellerangaben eingesetzt. In der Regel erfolgte während der Restriktion des Vektors ebenfalls eine Dephosphorylierung dessen, sowie im Anschluss eine Hitzeinaktivierung der Nukleasen. Sollte dieser Schritt nicht möglich sein, wurde das geschnittene Fragment aus dem Gel eluiert (3.2.4).

### 3.2.4 Gelextraktion von DNA

Die Extraktion von DNA Fragmenten aus Agarosegelen wird angewendet, sowohl bei einer PCR mit Nebenprodukten als auch bei einem verdauten Vektor, falls dieser oder das Insert weiterverwendet werden sollen.

Für die Gelextraktion wurde ein 0.3%-0.8%-iger Gel verwendet (3.2.2), nach dem Lauf wurde die Bande mittels UV-Licht ausgeschnitten, und das Gelstück mit dem Wizard® SV Gel and PCR Clean-Up System (Promega) oder Monarch Gel Extraction Kit (NEB) nach Herstellerangaben behandelt. Die Elution erfolgte in 15-30 µl ddH<sub>2</sub>O.

Die, aus diesem Schritt entstandenen gereinigten, DNA-Fragmente wurden anschließend mittels dem UV5-Spektrophotometer (Mettler Toledo) vermessen und in die Ligation (3.2.5) eingesetzt.

### 3.2.5 Ligation

Die Ligation wurde dazu genutzt, um Vektoren mit spezifischen DNA-Fragmenten zu erstellen. Sowohl der Vektor als auch das DNA-Fragment wurden mit Restriktionsenzymen behandelt, um spezifische 3'- und 5'-Überhänge zu generieren. Dabei wurde darauf geachtet, dass die Enden des Vektors mit denen des DNA-Fragments kompatibel sind. Die vorhandenen Lücken bei einer Ligation werden durch eine DNA-Ligase geschlossen.

Für den Ansatz wurde die T4-DNA-Ligase (Thermo Scientific) nach Herstellerangaben eingesetzt. Standardmäßig wurden 50-80 ng Vektor, bei einem 5-fachen Überschuss des DNA-Fragments verwendet, wobei die Menge an einzusetzen DNA-Fragment (Insert [ng]) mit folgender Formel berechnet wurde. Nach der Ligation wurde eine Transformation (3.1.4) durchgeführt.

$$\text{Insert [ng]} = \frac{\text{Überschuss x Vektor [ng]} \times \text{Insert [bp]}}{\text{Vektor [bp]}}$$

### 3.2.6 Plasmidisolation

Die Isolation der Plasmide wurde für verschiedene Zwecke durchgeführt. Zum einen dienten die Vektoren als Templat für PCR-Reaktionen, oder bei

Leervektoren als Grundvektor für Restriktionen. Zudem wurden die neu gebildeten Vektoren für eine Sequenzierung isoliert.

Für die eine kleine Präparation wurde das GeneJet Plasmid Mini Kit (Thermo Scientific), für größere entweder das Plasmid Midi Kit oder Plasmid Maxi Kit (Qiagen) verwendet.

### 3.2.7 Isolation von genomischer DNA

Für die Isolation von genomischer DNA (gDNA) wurde das QIAamp DNA Mini Kit (Qiagen) nach den Angaben für die Extraktion aus Gram-positive Zellen verwendet. Für das Protokoll wurden gefrorene Zellpellets von *M. thermautotrophicus* und *M. marburgensis* (AG Landau, MPI-Marburg) eingesetzt.

### 3.2.8 Klonierung der pET28a(+)-Vektoren

Der pET28a(+) ist ein high-copy Vektor, was bedeutet, dass er in 10-100 Kopien pro Zelle vorliegt. Er besitzt eine Resistenzkassette für das Antibiotikum Kanamycin. Durch eine geeignete Wahl der Restriktionsstellen kann an das entstehende Protein eine C- oder N-terminaler His-Fusion erfolgen.

Für die Klonierung wurde von folgenden Stämme die isolierte genomische DNA verwendet: *A. aeolicus* (AG Hartmann, Marburg), *A. ehrlichii* (DSMZ), *T. indicus* (DSMZ), *T. nitratireducens* (DSMZ), *H. halophila* (DSMZ), *M. thermautotrophicus* (3.2.7) und *M. marburgensis* (3.2.7). Für weitere Konstrukte wurde aus der NCBI Datenbank die entsprechende Gensequenz für die Gen-Synthese durch Biomatik (USA/Canada) verwendet, welche dann inklusive der Schnittstellen in pUC57 kloniert wurden: *H. volcanii*, *P. furiosus*, *M. infernorum* und *M. mazei*. Für die Klonierung der Gene aus *A. aeolicus* (aq\_221; aq\_707; aq\_880) wurde zunächst das PCR-Produkt in den Vektor pCR2.1TOPO (Invitrogen) eingebracht und sequenziert. Es folgte die Restriktion mit den Enzymen *NcoI* und *XhoI* (Fermentas) um das Insert zu erhalten. Für aq\_221 wurde das Insert durch die Enzyme *BspHI* und *HindIII* (Fermentas) generiert. Der Vektor pET28a(+) wurde mit den Enzymen *NcoI* und *XhoI* (Fermentas) beziehungsweise *NcoI* und *HindIII* (Fermentas) behandelt. Nach

einer Hitzeinaktivierung erfolgte die Ligation (3.2.5). Für die Klonierung der anderen Enzyme, mit Ausnahme von mlg\_2556 (*BspHI* und *XhoI*) wurde entweder das PCR-Produkt oder der pUC57-Klon (Tabelle 5), sowie der Leervektor pET28a(+) mit den Restriktionsenzymen *NcoI* und *XhoI* (ThermoScientific) behandelt. Im Falle des pUC57-Klons wurde das Insert durch eine Gelelution (3.2.4) gereinigt, wohingegen die anderen Reaktionen Hitzeinaktiviert wurden und für die Ligation eingesetzt werden konnten (3.2.5). Anschließend erfolgte die Transformation nach *E. coli* DH5 $\alpha$  (3.1.4).

*Tabelle 5: Auflistung der extern synthetisierten Plasmide für die Klonierung nach pET28a(+)*

Es sind sowohl der Organismus, sowie der Name des Gens, ob es Codon optimiert ist für *E. coli* und das erhaltene Plasmid angegeben.

Organismus	Gen	Codon optimiert	Plasmid
<i>H. volcanii</i>	hvo_1114	Ja	pUC57::hvo1114
<i>P. furiosus</i>	pfc_09185	Ja	pUC57::pfc09185
<i>M. infernorum</i>	minf_0007	Ja	pUC57::minf0007
<i>M. mazei</i>	mm_2077	Nein	pUC57::mm2077

### 3.2.9 Klonierung der pDG148(+) Vektoren

Der pDG148(+) ist ein high-copy Vektor, was bedeutet, dass er in 10-100 Kopien pro Zelle vorliegt. Er besitzt eine Resistenzkassette für das Antibiotikum Ampicillin. Dieser Vektor kann sowohl für *E. coli* als auch *B. subtilis* verwendet werden. Bei diesem System werden die zu untersuchenden Proteine hinter einen Spac-Promotor kloniert, welcher eine minimale Hintergrundexpression ohne Induktion zulässt. Bei Bedarf ist der Promotor durch IPTG induzierbar.

Dieses Vektorsystem wurde für die Untersuchung der *in vivo* Aktivität der gewünschten Proteine genutzt, welches sowohl mit als auch ohne C-terminale Histidin-Fusion erfolgte.

Für die Klonierung wurde mittels PCR (3.2.1) auf die vorhandenen Vektoren pET28a(+) oder pUC57 das gewünschte Insert amplifiziert. Sollte das Protein mit einer C-Terminale Histidin-Fusion versehen werden, wurde zusätzlich der T7-Terminator mit amplifiziert (

Tabelle 7, Primer 646). In der Regel wurde das Insert, sowie der Vektor pDG148(+) mit Hilfe der Restriktionsenzyme *SmaI* und *PaeI* verdaut (3.2.3), eine Ausnahme stellten die P Proteine da, welche mit *HindIII* und *PaeI* verdaut wurden. Nach einer Hitzeinaktivierung erfolgte die Ligation (3.2.5), sowie im Anschluss die Transformation nach *E. coli* DH5 $\alpha$  (3.1.4).

### 3.2.10 Klonierung der pGEX-mod Vektoren

Der pGEX-mod Vektor ist eine Form des pGEX-6P, welcher in der multiplen Klonierungsstelle eine *NdeI* Erkennungssequenz besitzt. Der Vektor wurde von der AG Reuter (Uni Marburg) bereitgestellt.

Die Klonierung in diesen Vektor sollte durch das Abspalten der Glutathione-Transferase-S (GST) Fusion ein annähernd natives Protein generieren. Dafür befindet sich zwischen dem GST-Tag und der Proteinsequenz eine Erkennungssequenz für die PreScission-Protease.

Für die Klonierung wurden die entsprechenden Gene durch eine PCR (3.2.1) auf schon vorhandene Vektoren mit den in Tabelle 8 aufgelisteten Oligonukleotiden durchgeführt. Das PCR Produkt wurde anschließend mit den Restriktionsenzymen *NdeI* und *BamHI* behandelt, sowie ebenfalls der Vektor pGEX-mod, welcher zusätzlich noch dephosphoryliert wurde. Durch eine Hitzeinkubation (80 °C, 10 min) wurden die Reaktionen abgestoppt und anschließend die Fragmente miteinander ligiert (3.2.5), transformiert (3.1.4) und durch eine Kolonie-PCR (3.1.6) analysiert.

*Tabelle 6: Auflistung der Oligonukleotide für die Klonierung in pET28a(+)*

Die verwendeten Oligonukleotide sind mit Nummer (Nr.) und Gen-Namen (Namen) aufgeführt. Schnittstellen sind fett hervorgehoben. aq = *A. aeolicus*, mlg = *A. ehrlichii*, thein = *T. indicus*, hhal = *H. halophila*, tvnir = *T. nitratireducens*, mtbma = *M. marburgensis*, mth = *M. thermotrophicus*, mm = *M. mazei*

Nr.	Name	Sequenz (5' -> 3')
630/ 631	aq_880	AAG <b>CCA TGG</b> ATG TGT TCG TTC TCG ACA C / CTC <b>TCG AGA</b> AAC CTG TGT CTT ACC AAG CTC TC
634 / 635	aq_707	AAG <b>CCA TGG</b> AAA TCC TTG TAA TAG GCG GTG G / CT <b>CTC GAG</b> CTC CTT GAA CCT GTA AGG GGA GAA AG



638 / 639	aq_221	AAG <b>TCA TGA</b> GTG TGG AGG AAA AAG CGA GGA TAG G / CTT <b>AAG CTT</b> GAC ATC GTC CGT GAA CTT GGG TCT TC
069 / 070	mlg_2556	TTT <b>TCA TGA</b> TGC GCC GAT TCG TGC / TTT <b>CTC GAG</b> ATG GCC GGC ACC CA
079 / 080	thein_2060	TTT <b>CCA TGG</b> ATG GCT TTT GAA TTA GAA GTT GAA AG / TTT <b>CTC GAG</b> AGG ACT GTA AAT AAG GCC TTC T
089 / 090	hhal_2443	TTT <b>CCA TGG</b> ATG CGC CGA TTC GTG C / TTT <b>CTC GAG</b> CCC GGC GGG CTG G
099 / 100	tvnir_2893	TTT <b>CCA TGG</b> ATG CGT CGC TTC GTT CTG G / TTT <b>CTC GAG</b> GCC GGA CGT CTT GGG
113 / 114	mtbma_C06930	TTT <b>CCA TGG</b> ATG CTT GCC AAG CAG AGA TTT G / TTT <b>CTC GAG</b> TCA TAA TAT TTA AGG TAT TCC TTG AGC ATC TTG GGG
121 / 122	mth_243	TTT <b>CCA TGG</b> ATG CTT GCA AAG CAG AGA TTC GTA C / TTT <b>CTC GAG</b> TTC ATA ATA TTT AAG GTA TTC CTT GAG CAT CTT GGG G
177a / 178	mm_2077	TTT <b>TTT TCA TGA</b> ATG AGA AAG TTT GTT TTA GAC ACC AGC / TTT <b>TCT CGA</b> GAG AAA CTG GCT CTC CGG T

*Tabelle 7: Auflistung der Oligonukleotide für die Klonierung in pDG148(+)*

Die verwendeten Oligonukleotide sind mit Nummer (Nr.) und Gen-Namen (Namen) aufgeführt. Schnittstellen sind fett hervorgehoben. aq = *A. aeolicus*, mlg = *A. ehrlichii*, thein = *T. indicus*, hhal = *H. halophila*, tvnir = *T. nitratireducens*, mtbma = *M. marburgensis*, mth = *M. thermautotrophicus*, mm = *M. mazei*, hvo = *H. volcanii*, minf = *M. inferorum*, pfc = *P. furiosus*

Nr.	Name	Sequenz (5' -> 3')
645 / 054a	aq_880	AGC <b>CCG GGC</b> ATG GAT GTG TTC GTT CTC GAC AC / TTT <b>GCA TGC</b> TTA AAA CCT GTG TCT TAC CAA GC

073 / 154	mlg_2556	TTT <b>CCC GGG</b> ATG CGC CGA TTC GTG C / TTT <b>CTC GAG</b> CTA ATG GCC GGC ACC CAG G
083 / 084	thein_2060	TTT <b>CCC GGG</b> ATG GCT TTT GAA TTA GAA GTT GAA AG / TTT <b>GCA TGC</b> TTA AGG ACT GTA AAT AAG GCC
088 / 093	hhal_2443	TTT <b>CCC GGG</b> ATG CGC CGA TTC GTG C / TTT <b>GCA TGC</b> CTA CCC GGC GGG CTG
094 / 097	tvnir_2893	TTT <b>TCT AGA</b> TGC GTC GCT TCG TTC TGG / TTT <b>GTC GAC</b> ATC CGG ATA TAG TTC CTC CTT TCA G
119 / 120	mtbma_C06930	TTT <b>CCC GGG</b> ATG CTT GCC AAG CAG AGA TTT G / TTT <b>TCT AGA</b> TCA TTC ATA ATA TTT AAG GTA TTC CTT GAG CAT CTT GGG G
127 / 128	mth_243	TTT <b>CCC GGG</b> ATG CTT GCA AAG CAG AGA TTC GTA C / TTT <b>TCT AGA</b> TCA TTC ATA ATA TTT AAG GTA TTC CTT GAG CAT CTT GG
175 / 176	hvo_1114	TTT <b>TTC CCG</b> GGA TGG CAG AAT ACC CGC TGA AAC / TTT <b>TGC ATG</b> CTT ACT CGA GAC GGT CAG GGT CC
173 / 174	mm_2077	TTT <b>TCC CGG</b> GAT GCT GAA GCA GCG CTT TG / TTT <b>TGC ATG</b> CTC ACT CGA GAA AAT CGG CAT CAT CGC
171 / 172a	pfc_09185	TTT TTC <b>CCGGGA</b> TGA TCC GTT TCA TCC TGG ATA CCA GC / TTT TTT <b>GCA TGC</b> TTA CTT ACG TTC CTT CTC ACT TTC CAC
646	ReverserPrimer für His <sub>6</sub> - Konstrukte	AGG <b>CAT GCA</b> TCC GGA TAT AGT TCC TCC TTT CAG
165a / 164	mlg_RS14575	TTT <b>AAG CTT</b> GTG GCG GCC CGA GG / TTT <b>GCA TGC</b> TTA GTG CCT GTT GCC GTC
165 / 166	hhal_1229	TTT <b>AAG CTT</b> TTG TCG GCC GAC GAG AGG / TTT <b>GCA TGC</b> TCA GGA GCC ATC GCA TGG
169 / 170	tvnir_3809	TTT <b>AAG CTT</b> GTG AGC GGG CCA GCT CC / TTT <b>GCA TGC</b> TCA TTT CAC GTG CTC CAG CAA C

217 / 168	thein_0634	TTT <b>CCC GGG</b> GTG GCG GCT TAC GGT ATA AAA AAA / TTT <b>GCA TGC</b> TCA ATA ACT TTT ATA ACT ATT TTT CTC ATA CTT TTT
500 / 501	minf_0924 (short)	GGA ATT AGC TTG <b>CAT GCT</b> CAT ATT TTA AGT TTA TCC CCC / TTT TTT TTT <b>CCC GGG</b> ATG GCT TTT AGA AAT CCG GC

*Tabelle 8: Auflistung der Oligonukleotide für die Klonierung in pGEX-mod*

Die verwendeten Oligonukleotide sind mit Nummer (Nr.) und Gen-Namen (Namen) aufgeführt. Schnittstellen sind fett hervorgehoben. *aq* = *A. aeolicus*, *mt* = *M. thermotrophicus*, *mm* = *M. mazei* und *pfc* = *P. furiosus*

Nr.	Name	Sequenz (5' -> 3')
238 / 239	aq_88 0	TTTTTTTTTT <b>CATATGG</b> ATGTGTTCTCGACACG / TTTTTTTTTT <b>GG ATC</b> CTTAAACCTGTGTCTTACCAAGCTCTC
240 / 241	mm_2 077	TTTTTTTTTT <b>CAT ATG</b> CTGAAGCAGCGCTTTGTTCT / TTTTTTTTTT <b>GG ATC</b> CTCACCTCGAGAAAATCGGCATCATCGC
244 / 245	mt_2 43	TTTTTTTTTT <b>CA TAT G</b> CTTGCAAAGCAGAGATTCGTACTT / TTTTTTTTTT <b>GGA</b> <b>TCCTCATT</b> CATAATATTTAAGGTATTCCTTGAGCATCTT
248 / 249	pfc_09 185	TTT TTTTTC <b>TCA TAT GAT</b> CCG TTT CAT CCT GGA TAC CAG C / TTTTTTTTTTGGATCCTTACTTACGTTTCCTTCTCACTTTCCACTT TTTC

### 3.2.11 *In vitro* Mutagenese von Aq880 und Mth243

Die *in vitro* Mutagenese wurde durchgeführt, um einzelne Nukleotide aus zu tauschen, um auf Proteinebene einen Aminosäureaustausch hervorzurufen.

Das Protokoll wurde in Anlehnung an QuikChange® XL Site-Directed Mutagenesis Kit (Agilent Technologies) verändert. Zum einen wurden die Polymerasen *peqGOLD Pwo* DNA Polymerase (Pqlab) oder *Platinum SuperFi* DNA Polymerase (Invitrogen) verwendet, sowie zusätzlich komplementäre (Art 1) oder Rücken-an-Rücken (Art 2) Primer. Dabei wurde darauf geachtet, dass sich die zu tauschende Base mittig in dem Mutagenese-

Primern befindet. Eine Auflistung der Oligonukleotide für die Mutation von aq\_880 befindet sich in Tabelle 9, sowie für mth\_243 in

Tabelle 10. Es wird zusätzlich die Anlagerungstemperatur ( $T_A$ ), sowie die geänderte Base (**fett**) und das geänderte Codon (unterstrichen) mit angeben. Um welche Art der Mutagenese es sich handelte ist ebenfalls aufgeführt.

Die PCR (3.2.1) wurde nach Herstellerangaben durchgeführt und durch Agarose Gelelektrophorese (3.2.2) kontrolliert. Im Falle von komplementären Primern wurden 20 µL des PCR-Produkts mit 2 µL *DpnI* (Fermentas) für 3 h bei 37 °C behandelt. Im Anschluss wurden 4 µl des *DpnI*-in-vitro-Mutagenese Ansatzes für die Transformation nach *E. coli* DH5α (3.1.4) genutzt. Wurden die Rücken-an-Rücken Primer genutzt, erfolgte nach der Kontrolle des PCR Produkts eine Phosphorylierung des 5'-OH Ende mit Hilfe der T4-Polynukleotid-Kinase (T4 PNK, Thermo Scientific) nach Herstellerangaben, wobei 10 µl des PCR Ansatzes verwendet wurden. Anschließend erfolgte eine Kombination aus Ligation und Restriktion. Dabei wurden 10 µl des phosphorylierten PCR-Produkts für eine Ligation (3.2.5) eingesetzt mit zusätzlich 1 µL *DpnI* (Thermo Scientific) im Ansatz. Nach der Inkubation erfolgte die Transformation nach *E. coli* DH5α (3.1.4).

*Tabelle 9: Auflistung der Primer für die in vitro Mutagenese von aq\_880*

Die verwendeten Primer sind mit der entsprechenden Primernummer (Nr.), sowie der Anlagerungstemperatur ( $T_A$ ) und dem Namen angeben. Die Sequenz ist in 5' ->3' dargestellt, wobei die geänderte Base fett gedruckt und das geänderte Codon unterstrichen ist. Bei der Art handelt es sich bei 1 um komplementäre Primer (es ist nur einer der beiden angeben) und bei 2 um Rücken an Rücken Primer (es sind beide Primer angegeben, getrennt durch einen Querstrich).

aq_880				
Nr.	Name	Sequenz (5'->3')	$T_A$	Art
028/ 029	D7A	ATG TGT TCG TTC TCG <u>CCA</u> CGA GCG TTT TCA CAA ACC CGG	60°C	1
024/ 025	D138A	TCA GGC AGG GAA TAC TCG <u>CCA</u> GTA AAG AGG ACG TGG	60°C	1
009/ 010	D142A	AAT ACT CGA CAG TAA AGA <u>GGC</u> CGT GGA TGT GCT GCT CCT TGC C	60°C	1
026/ 027	D144A	AAT ACT CGA CAG TAA AGA GGA CGT <u>GGC</u> TGT GCT GCT CCT TGC C	60°C	1

013/ 014	D142A D144A	AAT ACT CGA CAG TAA AGA <u>GGC</u> CGT <u>GGC</u> TGT GCT GCT CCT TGC C	60°C	1
011/ 012	D160A	ACG GGG TTC TCG TTT CGG <u>CGG</u> CTG AAG GCC TCA GAA CAT GGG C	60°C	1
036/ 037	E141A	GGG AAT ACT CGA CAG TAA <u>AGC</u> GGA CGT GGA TGT GCT GCT	60°C	1
034/ 035	E141D	GGG AAT ACT CGA CAG TAA <u>AGA</u> CGA CGT GGA TGT GCT GCT	60°C	1
179/ 180	$\Delta 2$ -20	CAT GGT ATA TCT CCT TCT TAA AGT TAA AC/ GAA GAG GAT CAG AGA GGA GCT ATG GAA A	63°C	2
179/ 181	$\Delta 2$ -50	CAT GGT ATA TCT CCT TCT TAA AGT TAA AC/ GAAATGAGAAAAATAATGGATGTGGGGG	63°C	2
201/ 202	L93P	CAAACCTTTATTAATCCTATATCT <u>GGG</u> TTCTTCTA TGAACATAGAG	60°C	1
203/ 205	L124P	GCCCTCAGGCAGGGAATACT/ TTCCCTGTACTTCTCCCT <u>GGG</u> CCTAGCTATTAAT TTCCCGAC	60°C	2

*Tabelle 10: Auflistung der Primer für die in vitro Mutagenese von mth\_243*

Die verwendeten Primer sind mit der entsprechenden Primernummer (Nr.), sowie der Anlagerungstemperatur ( $T_A$ ) und dem Namen angeben. Die Sequenz ist in 5' ->3' dargestellt, wobei die geänderte Base fett gedruckt und das geänderte Codon unterstrichen ist. Bei der Art handelt es sich bei 1 um komplementäre Primer (es ist nur einer der beiden angegeben) und bei 2 um Rücken an Rücken Primer (es sind beide Primer angegeben, getrennt durch einen Querstrich).

mth_243				
Nr.	Name	Sequenz (5' ->3')	$T_A$	Art
199/ 214	D166A	TGA TCT GTA TTT CTT CCT GAG GTC CTT TAT / CCCTCCGCAAGGGAACCCTTG <u>CC</u> AGTGCCCCCGACCTT	60°C	2
195/ 213	D170A	AAG GGT TCC CTT GCG GAG GGC/ GACAGTGCCCCCG <u>CC</u> CTTGATGTGCTGCTCCTT GCCAAGGAA	60°C	2
197/ 213	D172A	AAG GGT TCC CTT GCG GAG GGC/ GACAGTGCCCCCGACCTTG <u>CT</u> GTGCTGCTCCTT GCCAAGGAA	60°C	2

193/ 212	D188A	CCT TGG CAA GGA GCA GCA C / AACTGGGGGCAGGGGTCGTGGCGGCC <u>CT</u> GAT GGTATAAGGGTCTGGG	60°C	2
209/ 211	L152P	AAGCACCTCGCTCTCGATCTC/ GGAAAGGCCATAAAGGAC <u>CC</u> CAGGAAGAAATAC AGATCA	60°C	2

### 3.2.12 Sequenzierung der Vektoren

Die Sequenzierung der Vektoren wurde durchgeführt, um eine erfolgreiche Klonierung oder *in vitro* Mutagenese zu bestätigen. Diese wurde von den Firmen MWG Eurofins (Deutschland) oder SeqLab (Deutschland) mittels Sanger-Sequencing durchgeführt. Die Ergebnisse wurden als .FASTA Datei gesendet und mittels der Programme SerialCloner 2.6 (Freeware, [http://serialbasics.free.fr/Serial\\_Cloner.html](http://serialbasics.free.fr/Serial_Cloner.html)) oder SnapGene-Software (GSL Biotech; Version 4, [www.snapgene.com](http://www.snapgene.com)) mit der gewünschten DNA Sequenz abgeglichen.

### 3.3 Herstellung von verschiedenen Prä-tRNA Substraten

Für die Charakterisierung der in dieser Arbeit genutzten Proteine wurden verschiedene Prä-tRNA Substrate untersucht. Dabei wurde als Standard-Substrat die Prä-tRNA<sup>Gly</sup> von *Thermus thermophilus* verwendet, welche einen 14 nt-langen 5'-Überhang und 6 nt am 3'-Ende (inklusive CCA) besitzt.

Für die Herstellung dieser Substrate wurden zunächst die entsprechenden Plasmide isoliert (3.2.6) und linearisiert (3.2.3). Anschließend wurde eine T7-transkription (3.3.5) mit dem linearisierten Plasmid durchgeführt und die RNA Geleluiert (3.3.7). Zusätzlich zu den Plasmiden wurden ebenfalls PCR-Templates eingesetzt, welchen entweder ein Plasmid oder ein Templat-Oligo als Ausgangsmaterial zu Grunde lag. Eine genaue Auflistung, sowie die Herstellung ist in Brillante *et al.* zu finden [108].

### 3.3.1 Vektoren mit diversen Varianten der Prä-tRNA<sup>Gly</sup>

In dieser Arbeit wurden Plasmide verwendet, welche unterschiedliche Längen der 5'-Flanke der Prä-tRNA<sup>Gly</sup> aufwiesen, sowie Variationen des Akzeptorstamms und Deletionen der tRNA Domänen. Eine Auflistung der verwendeten Plasmide, sowie das verwendete Restriktionsenzym zur Linearisierung des Plasmides sind in Tabelle 11 aufgelistet.

*Tabelle 11: Auflistung der Plasmide für 5'-Längen Varianten von Prä-tRNA<sup>Gly</sup>*

Auflistung der Plasmide, welche als Templat dienten um verschieden 5'-Varianten Prä-tRNA<sup>Gly</sup> zu erhalten. Es ist sowohl die Bezeichnung des Substrats (Prä-tRNA<sup>Gly</sup>), als auch das Restriktionsenzym für die Linearisierung angegeben.

Prä-tRNA <sup>Gly</sup>	Vektor	Linearisierung
L14	pSBpt3-hh-L14	<i>BamHI</i>
L7	pSBpt3-hh-L7	<i>BamHI</i>
L4	pSBpt3-hh-L4	<i>BamHI</i>
L2	pSBpt3-hh-L2	<i>BamHI</i>
L1	pSBpt3-hh-L1	<i>BamHI</i>
L14 Aa+2	pTtGly-Aa+2	<i>BamHI</i>
L14 Aa-2	pTtGly-Aa-2	<i>BamHI</i>
AaT	pAaT-TtGly	<i>BamHI</i>
Aab9T	pAab9T-TtGly	<i>BamHI</i>

### 3.3.2 Herstellung der Template für die T7 Transkription

Um genug Plasmid zu erhalten, wurden die Plasmide durch eine mittlere Präparation (siehe Plasmidisolation 3.2.6) isoliert. Anschließend wurden die Plasmide mit Hilfe des Enzyms *BamHI* (Fermentas) linearisiert. Dazu wurden 20-40 µg Plasmid-DNA in ein maximales Ansatzvolumen von 50 µl verdaut. Als Kontrolle wurden 300 ng des verdauten Vektors auf ein 0.8%-iges Agarosegel (3.2.2) aufgetragen und analysiert. Insofern das Plasmid vollständig linear vorlag, wurde eine Phenol-Chloroform-Reinigung (3.3.3) durchgeführt.

Zur Herstellung der Minimalsubstrate wurde eine PCR durchgeführt (3.2.1) mit den in Tabelle 12 aufgeführten Oligonukleotiden, anschließend erfolgte die Reinigung, entweder durch ein Agarose-Gel oder nur der PCR-Ansatz (3.2.4). Die Elution des Templat erfolgte in 40  $\mu$ L Nuklease-freies Wasser und wurde vollständig für die Reaktion (3.3.5) genutzt.

*Tabelle 12: Auflistung der Oligonukleotide für die Herstellung von Minimalsubstraten*

Aufgeführt sind die verschiedenen Oligonukleotide, welche zur Herstellung der Minimalsubstrate verwendet wurden. For: Primer 1, Rev: Primer 2. Bei fehlendem Templat wurde der Vektor pSBpt3-hh-L14 verwendet. Ausnahme Aa<sub>b9</sub>T<sub>mut</sub>, da es sich um teilweise Komplementäre Primer handelt, wurde kein Templat benötigt. Eine genaue Auflistung der Bedingungen findet sich in der Referenzpublikation.

Minimalsubstrate		
Name	Sequenz (5' ->3')	Ref.
$\Delta$ Ac	<b>For:</b> ATATATATATATAATACGACTCACTATAGG <b>Templat:</b> TAATACGACTCACTATAGGATTTTCCCTTTTCGCGGGAGTAGCTCAGTCGGTAG AGCACCTTCGGGGGGTCGCGGGTTCAAGTCCCGTCTCCGCTCCAATA <b>Rev:</b> TATTGGAGCGGGAGACG	[108]
$\Delta$ D	<b>For:</b> ATATATATATATAATACGACTCACTATAGG <b>Templat:</b> TAATACGACTCACTATAGGATTTTCCCTTTTCGCGGGAGTAGACCGACCTTGCC AAGGTCGGGTCGCGGGTTCAAGTCCCGTCTCCGCTCCAATA <b>Rev:</b> TATTGGAGCGGGAGACG	
Aa <sub>b9</sub> T <sub>mut</sub>	<b>For:</b> CTCGAGTAATACGACTCACTATAGGATTTTCCCTTTTCGCGG GAGCCAGCTCCTG <b>Rev:</b> GACTGGAGCGGGAGGCGGGTAAAAGGCCCG CAGGAGCTGGC	
CCA	<b>For:</b> ATATATATATATAATACGACTCACTATAGG <b>Rev:</b> TGGAGCGGGAGACGGGACTT	
$\Delta$ CCA	<b>For:</b> ATATATATATATAATACGACTCACTATAGG <b>Rev:</b> TATTGGAGCGGGAGACG	
40 nt	<b>For:</b> ATATATATATATAATACGACTCACTATAGG <b>Rev:</b> TCCTCACGGACTCATCAG	

### 3.3.3 Phenol-Chloroform-Reinigung

Die Phenol-Chloroform-Reinigung wird angewendet um Nukleinsäure-haltige wässrige Lösungen effizient von Proteinen abzutrennen. Diese verbleiben, denaturiert in der organischen Phase, während die Nukleinsäuren in der wässrigen Phase verbleiben [117].



Handelte es sich bei dem zu reinigenden Ansatz um das Templat für eine T7-Transkription wurde das Volumen mit ddH<sub>2</sub>O auf 200 µl aufgefüllt. Anschließend wurde das gleiche Volumen an Phenol (DNA) beziehungsweise saurem Phenol (RNA) dazu gegeben. Dieser Ansatz wurde für 1,5 min gevortext und für 5 min bei 10 000 rpm zentrifugiert. Die dadurch entstandene obere wässrige Phase wurde vorsichtig in ein neues Reagiergefäß überführt und mit dem gleichen Volumen an Chloroform versetzt, ebenfalls 1,5 min gevortext und 5 min bei 10 000 rpm zentrifugiert. Zum Schluss wurde die obere wässrige Phase erneut in ein neues Reagiergefäß überführt und für die Ethanol-Fällung (3.3.4) genutzt.

### 3.3.4 Ethanol-Fällung

Die Ethanol-Fällung wird vor allem angewendet um die Nukleinsäuren zu konzentrieren und Konzentrationen an Salzen zu minimieren [118].

Die durch die Phenol-Chloroform-Fällung erhaltene wässrige Lösung wurde mit  $\frac{1}{10}$  des Volumens an 3 M Natriumacetat pH 5 und dem 2,5-fachen Volumen an reinem Ethanol versetzt. Der Ansatz wurde gut gemischt und entweder bei -80°C für mindestens 1 h oder über Nacht bei -20°C gelagert. Anschließend erfolgte eine Zentrifugation bei maximaler Umdrehung und 4°C für mindestens 1 h. Der Überstand wurde verworfen und das Pellet für 5 - 10 min bei Raumtemperatur getrocknet. Zum Schluss wurde das Pellet in Nuklease-freiem Wasser aufgenommen, die Konzentration bestimmt und bei -20°C bis zur Verwendung gelagert.

### 3.3.5 T7 *in vitro* Transkription

Die Isolierung von T7 RNAP erfolgt aus dem *E. coli*-Stamm BL21 pY639F. Die T7 RNAP besitzt einen Aminosäureaustausch an Position 693 von Tyrosin (Y)- gegen Phenylalanin (F), dadurch können neben normalen Ribonukleotiden (rNTPs) auch NTPs mit 2'-Ribose-Modifikationen (z.B. 2'-Fluoro-NTPs, 2'-Amino-NTPs, 2'-Deoxy-NTPs) als Substrat verwendet und somit in den RNA-Strang eingebaut werden.

Für die T7-Transkription wurde zunächst ein analytischer Ansatz (100 µL) durchgeführt, bevor der präparative (1 - 5 mL) angesetzt wurde. Die Zusammensetzung des Reaktionsansatzes findet sich in Tabelle 13. Während eines analytischen Ansatzes wurden nach 2 h und 4 h je 4 µl Probe zu 4 µl 2x PPF gegeben und auf einem analytischen 8%-igen denat. PAA-Gel ausgewertet (3.3.6). Nach einem präparativen Ansatz erfolgte zunächst eine erneute Zugabe an T7 RNA Polymerase nach 2 h und nach Ablauf der 4 h Inkubation bei 37°C eine Phenol-Chloroform Reinigung (3.3.3) mit anschließender Ethanol-Fällung (3.3.4). Das gelöste Produkt der Ethanol-Fällung wurde mit dem gleichen Volumen 2x PPF versetzt und auf ein präparatives 8%-iges denat. PAA-Gel aufgetragen (3.3.6) und anschließend Geleluert (3.3.7).

*Tabelle 13: Ansatz einer T7 in vitro Transkription*

Aufgelistet sind die Endkonzentrationen eines präparativen Ansatzes. Die Inkubation erfolgte bei 37°C für 4 h, wobei nach 2 h noch einmal T7 RNA Polymerase zugeben wurden.

<b>T7 in vitro Transkription</b>	
20 mM	MgCl <sub>2</sub>
80 mM	Hepes, pH 7,5
1 mM	Spermidine
5 mM	DTT
0.05 U	Pyrophosphatase
40 µg	Templat, linearisiert
3.75 mM	Je NTP
5 %	DMSO
<sup>1</sup> / <sub>50</sub> V	T7 RNA Polymerase

### 3.3.6 Denaturierende Polyacrylamid-Gelelektrophorese (PAGE)

Die Polymerisationsreaktion von Polyacrylamid (PAA) wird durch die Zugabe von <sup>1</sup>/<sub>1000</sub> V N,N,N',N'-Tetramethylethyldiamin (TEMED) und <sup>1</sup>/<sub>100</sub> V 10% (w/v) Ammoniumperoxodisulfat (APS) gestartet. Dabei ist der Grad der Quervernetzung abhängig vom Acrylamid:Bisacrylamid Verhältnis, wobei die

Porengröße durch die Acrylamid-Konzentration bestimmt wird. Je nach Größe der zu untersuchenden Fragmente wurde die Konzentration angepasst, ausgehend von einer 20% PAA Stammlösung (8 M Urea, 20% PAA (50%, 48:2), 1x TBE (89 mM Tris, 89 mM Borsäure, 2 mM EDTA), ad 1 L ddH<sub>2</sub>O).

Die mit 1x Ladepuffer (2x; 0,02% (w/v) Bromphenolblau, 0,02% (w/v) Xylencyanoblau, 2,6 M Urea, 66% (v/v) Formamid in 2x TBE) versetzten Proben wurden in die entsprechenden Proben-Taschen gegeben und das Gel in 1x TBE gefahren (10 - 30 mA). Anhand des Migrationsverhalten der verwendeten Farbstoffe wurde bei gewünschter Trennleistung das Gel gestoppt und entweder gefärbt (Ethidiumbromid, GelRed) oder mittels Autoradiographie untersucht.

Für Aktivitätsanalysen wurden 20 - 25%-ige Gele verwendet und für die Reinigung und Analyse der T7-Transkripte, sowie das Radiolabeling 8%.

### 3.3.7 Reinigung der T7-Transkription mittels Gel Elution

Die Reinigung mittels einer Gel Elution erfolgte zum einen um die Enzyme und Puffer abzutrennen, sowie die Auftrennung der unterschiedlichen langen Transkripte.

Für die Reinigung der T7-Transkripte wurde ein 8%-iges denat. PAA-Gel verwendet mit einer einzelnen großen Tasche. Der zuvor mit Phenol und Chloroform gereinigte Transkriptionsansatz wurde nach einer Ethanol Fällung in ddH<sub>2</sub>O aufgenommen und mit dem gleichen Volumen 2x PPF versetzt und anschließend auf das Gel aufgetragen. Da die meisten, in dieser Arbeit verwendeten Transkripte mit der oberen Farbstoffbande migrierten, wurde diese bis zur Mitte des Gels wandern gelassen, bevor das Gel gestoppt wurde.

Nach dem Lauf wurde das Gel mittels UV *shattering* bei 580 nm behandelt und die Transkript Bande markiert. Diese wurde mit einem Skalpell ausgeschnitten und in 1 M NaOAc (5 - 10 mL) gegeben, über Nacht bei 4°C geschüttelt und am nächsten Tag mit dem 2,5 x Volumen an Ethanol p.a versetzt. Der Fällungsansatz wurde für 1 h bei -80°C gelagert, für 1 h bei 8 000 rpm zentrifugiert (4°C) und der Überstand anschließend verworfen. Nachdem das Pellet getrocknet war, wurden 50 - 100 µl ddH<sub>2</sub>O genutzt, um das Pellet zu lösen. Das gereinigte Transkript wurde bei -20°C bis zur Verwendung gelagert.

### 3.3.8 Radioaktive Markierung der Substrate am 5'-Ende

Für die radioaktive Markierung der Substrate am 5'-Ende wurde gamma-ATP (32-P) eingesetzt.

Zunächst wurden 30 pmol an Substrat am 5'-Ende dephosphoryliert mit Hilfe einer alkalinen Phosphatase (Fast-AP, Thermo Scientific) nach Herstellerangaben mit anschließender Hitzeinaktivierung. Für die Markierung wurden ein 10 µL Ansatz genutzt, welcher sich nach den Herstellerangaben der T4-PNK (Thermo Scientific) richtete, wobei statt ATP insgesamt 3 µL gamma-ATP 32-P zugesetzt wurden. Dieser Ansatz wurde für 1 h bei 37°C inkubiert und anschließend auf einem 8% denat. PAA-Gel aufgetrennt. Der Erfolg der Markierung wurde durch das Auflegen einer Imaging Plate für 60 - 90 Sekunden mit anschließender Auswertung durch einen Phosphoimager (Bio-Imaging Analyser BAS 1000, Fujifilm) überprüft. Das markierte Substrat wurde aus dem Gel ausgeschnitten und über Nacht bei 4°C in 500 µL 1 M NaOAc eluiert, am nächsten Tag mit dem 2,5x Volumen an 100% Ethanol versetzt und für 3 - 6 h bei -20°C gefällt. Nach einer einstündigen Zentrifugation (13 000 rpm, 4 °C) wurde das getrocknete Pellet in 10 µL ddH<sub>2</sub>O gelöst und letztendlich mit dem TriCarb (Perkin Elmer) vermessen. Die Lagerung des radioaktiven Substrates erfolgte in einem Bleigefäß bei -20°C.

## 3.4 *In vitro* Analysen von rekombinanten Proteinen

Die in dieser Arbeit verwendeten rekombinanten Proteine wurden mit Hilfe einer terminalen Histidin-Wiederholung exprimiert. Da die Proteine, welche analysiert werden sollten, überwiegend aus extremophilen Organismen stammten, wurde bei der Expression darauf geachtet ein Expression System zu nutzen, welche eine optimale Ausbeute ermöglichte. In Tabelle 4 sind die entsprechenden Proteine (durch ihre Genbezeichnung) aufgelistet.

### 3.4.1 Expression der rekombinanten Proteine

Für die Expression der Proteine wurde der Stamm *E. coli* Rosetta(DE3), *E. coli* BL21 (DE3) oder *E. coli* BL21 (DE3) pLysS verwendet. Der Stamm Rosetta besitzt das pRARE Plasmid, welches seltene Codons der tRNAs trägt. Dadurch

können Proteine trotz einer nicht-optimalen Codon Nutzung in *E. coli* mit einer verbesserten Ausbeute exprimiert werden.

Die Induktion der spezifischen Proteinexpression in *E. coli* Rosetta (DE3) wurde durch 0,2% Laktose in einem selbst-induzierendem LB-Medium (LB-AIM, 0,05% Glukose, 0,2% Laktose, verändert nach Studier [119]) bei 37°C und 180 - 220 rpm durchgeführt. Es wurde eine 800 - 1000 mL Kultur in Schikanen Kolben angelegt, welche entweder durch eine ausgestrichene Einzelkolonie oder direkt aus der Glycerolkultur angeimpft wurde. Die Expression wurde über Nacht durchgeführt und die Zellen am nächsten Tag geerntet (4°C, 3 000 rpm). Die Zellen wurden bei Bedarf bei -80°C bis zur Reinigung der Proteine gelagert.

Die Induktion von den *E. coli* BL21-Stämmen erfolgte mit 1 mM IPTG. Dafür wurde eine Übernachtskultur 1:100 - 1:1000 in frisches LB Medium verdünnt und das Wachstum photometrisch überwacht. Bei einer OD<sub>600</sub> von 0.3 - 0.6 wurde das IPTG zugegen und die Zellen nach 4 h geerntet (4°C, 3 000 rpm). Zusätzlich konnte dieser Stamm auch mit Hilfe der Autoinduktion verwendet werden. Die Zellen wurden bei Bedarf bei -80°C bis zur Reinigung der Proteine gelagert.

### 3.4.2 Proteinreinigung mittels des Batchverfahrens

Die Affinitätsreinigung mittels des Batch-Verfahrens wurde für erste Analysen des entsprechenden Proteins genutzt. Für diese Methode wurde Protino®-Ni-NTA (Macherey-Nagel) eine Matrix, welche mit Nickelionen beladen ist, genutzt.

Die (aufgetauten) Zellen wurden in 15 mL NPI-20 gelöst und mit Hilfe von Ultraschall aufgeschlossen. Dabei wurden drei Zyklen mit einer Intensität von 50% und einem Output von 5 gefahren. Zwischen den Durchgängen wurden die Zellen für 3 min auf Eis gelagert. Anschließend wurden die Zelltrümmer durch eine Zentrifugation (4°C, 8 000 rpm) von dem Protein-haltigen Lysat abgetrennt. Währenddessen wurde die Matrix vorbereitet. Zunächst wurden 2 mL des Gemisches, was 1 mL reiner Matrix entspricht, zentrifugiert (4°C, 3 000 rpm) und der Überstand abgenommen. Im Anschluss wurde die Matrix zweimal mit 2 mL NPI-20 gewaschen und mit 5 mL NPI-20 equilibriert. Für die Bindung der Proteine an die Matrix wurde der Überstand für 45 - 60 min bei 4°C mit selbiger rotierend inkubiert. Um nicht gebundene Proteine zu entfernen wurde 3 - 5-mal mit NPI-50

gewaschen (invertiert, 4°C, 3 000 rpm). Zum Schluss wurde das Protein mit 1-2 mL NPI-250 (invertiert, 4°C, 3 000 rpm) eluiert. Das Eluat wurde über Nacht gegen Storage oder EMSA Puffer (Tabelle 14) dialysiert und bei -20°C gelagert.

### 3.4.3 Dialyse von gereinigten Proteinen

Die Dialyse der Proteine diente zum Austausch der Elutionspuffer der Reinigung gegen die entsprechenden Lagerpuffer. Dabei unterschieden sich die Lagerpuffer je nach Art der Anwendung und Experimente der Proteine hinterher. In Tabelle 14 sind alle in dieser Arbeit verwendeten Lagerpuffer aufgeführt, inklusive der Verwendung des Proteins im Anschluss.

Für die Dialyse wurde ein Dialyseschlauch mit einer Ausschlussgröße von 6-8 kDa (ZelluTrans/Roth Dialysiermembranen T2, Carl Roth) verwendet. Dieser wurde kurz in der Mikrowelle aufgekocht und anschließend etwas abgekühlt. Der Schlauch wurde durch Schlauchklammern verschlossen. Die Dialyse erfolgte mit einem mindestens 100-fachen Überschuss des Lagerpuffers im Verhältnis zum Volumen der Proteinfractionen unter leichtem Rühren bei 4°C über Nacht. Nach einer Stunde wurde der Lagerpuffer noch einmal gewechselt. Am nächsten Tag wurde das Protein aus dem Dialyseschlauch entnommen und gelagert.

*Tabelle 14: Übersicht der Lagerpuffer für Proteine*

In dieser Tabelle sind alle Lagerpuffer für Proteine aufgelistet, inklusive ihrer Zusammensetzung und der späteren Verwendung der Proteine

Lagerpuffer	Zusammensetzung	Verwendung
Storage-Buffer	10 mM Tris pH 8 10 mM MgCl <sub>2</sub> 100 mM KCl 0,1 mM EDTA 3 mM DTT 20-50% Glycerin	Aktivitätsanalysen Oligomerisierungsanalysen Größenausschlusschromatographie Ionenaustauschchromatographie
EMSA-Puffer	10 mM Tris pH 8 10 mM CaCl <sub>2</sub> 100 mM KCl 0,1 mM EDTA 3 mM DTT 20-50% Glycerin	RNA-Bindungs-Analysen

Kristallisations- Start-Puffer	10 mM Tris pH 8 10 mM MgCl <sub>2</sub> 100 mM KCl 0,1 mM EDTA 3 mM DTT	Ionenaustauschchromatographie
Kristallisations- Puffer	20 mM Tris pH 8 100 mM KCl 0,1 mM EDTA 3 mM DTT	Strukturaufklärung durch Kristallisation und NMR
T7-RNAP- Lagerpuffer	20 mM Na-Phosphat pH 7,7 1 mM Na <sub>2</sub> EDTA 1 mM DTT 50% Glycerin (v/v) 2 µg/mL PMSF 10 mM NaCl	T7 <i>in vitro</i> Transkription
CD-Puffer	50 mM Na <sub>2</sub> HPO <sub>4</sub> 50 mM NaCl	CD-Spektroskopie

#### 3.4.4 Proteinreinigung für die Kristallisation

Einige der in dieser Arbeit verwendeten Proteine sollten für eine Strukturaufklärung mittels Kristallisation gereinigt werden. Dabei wurden 1-2 L Kulturvolumen angesetzt (3.4.1). Anschließend erfolgte eine erste Reinigung mittels der Affinitätschromatographie über eine FPLC (3.6.1). Nach einer ersten Analyse über ein SDS-PAGE (3.7.6) mit anschließender Coomassie-Färbung (3.7.7), wurde das Protein über Nacht in den Kristallisations-Start-Puffer (Tabelle 14) für die Ionenaustauschchromatographie dialysiert und anschließend mit einem starken Ionentauscher (MonoQ™ 5/50 GL, GE Healthcare) weiter gereinigt (3.6.2). Die ausgewählten Fraktionen wurden in den Kristallisations-Puffer (Tabelle 14) über Nacht dialysiert, die Konzentration mit Hilfe eines Bradfords bestimmt, in 200 µL Aliquot schockgefroren und an die AG Reuter (FB16, Uni Marburg) ausgehändigt. Zusätzlich wurde das Protein Aq880cHis, sowie Hvo1114cHis an MarXtal (Uni Marburg) für einen Roboter-basierten Screen gesendet und nach einer erfolgreichen Kristallbildung die Bedingungen in Kooperation mit der AG Reuter bestätigt, sowie die Kristalle weiter optimiert.

### 3.4.5 Proteinreinigung der T7 RNA Polymerase

Die in dieser Arbeit gereinigte T7 RNA Polymerase wurde für die Herstellung der verschiedenen RNA-Substrate verwendet (3.3.5). Die Reinigung erfolgt in Anlehnung an die Protokolle von Grudberg und Gunn [120], sowie Zawadzki und Gross [121] mit anschließender Ionenaustauschchromatographie (3.6.3). Die Herstellung der T7 RNA Polymerase wurde durch Studenten in einem Praktikum durchgeführt.

### 3.4.6 Bestimmung der Proteinkonzentration mittels Bradford

Die Bestimmung der Proteinkonzentration mittels einem Bradford wurde 1976 von Marion Bradford publiziert [122]. Das Prinzip der Bestimmung beruht auf einem colometrischen Umsatz von Coomassie. Bei einer Kontrolle ohne Protein (roter Farbstoff) hat die Lösung ein Absorptionsmaximum von 470 nm, sobald durch Proteine eine Komplexbildung eintritt (blaue Färbung) verschiebt sich das Maximum auf 595 nm.

Für die Bestimmung der Proteinkonzentration wird zuerst eine Kalibrierkurve mittels BSA angelegt. Dabei reicht die Proteinkonzentration von 0-20 µg/µL. Die Kalibrierung wird wie folgt durchgeführt, zuerst wird das entsprechende Volumen an BSA-Lösung vorgelegt, das Volumen mittels ddH<sub>2</sub>O auf 800 µL aufgefüllt und anschließend mit 200 µL Bradford-Reagenz (BioRad) versetzt. Es folgt eine Inkubation von 5 min bei Raumtemperatur bevor die Absorption bei 595 nm bestimmt wird. Die gemessenen Absorptionen werden gegen die Proteinkonzentrationen aufgetragen, so dass eine Kalibriergerade entsteht. Anschließend werden die gewünschten Proteine in zwei bis drei unterschiedlichen Volumina nach dem gleichen Protokoll photometrisch gemessen. Es wurden zwei unterschiedliche Photometer in dieser Arbeit verwendet, welche ein voreingestelltes Bradford Programm enthielten, so dass automatisch die Berechnung der Kalibrierkurve, sowie die Konzentration des Proteins ermittelt wurde. Die erhaltende Proteinkonzentration wurde, falls notwendig mit dem Verdünnungsfaktor (maximal 1:10) multipliziert und anschließend durch das eingesetzte Volumen (1-10 µL) dividiert. Durch diese Berechnung erhielt man eine Konzentration in µg/µL. Zum Schluss wurden die



zwei bis drei unterschiedliche Messungen mit unterschiedlichen Volumen addiert und gemittelt. Mit dieser Konzentration wurde anschließend die Molare Masse des Proteins berechnet.

#### **3.4.7 Aktivitätsanalyse der Proteine (nicht-radioaktiv)**

Die Aktivitätsanalyse der Proteine mit nicht radioaktiv markiertem Substrat wurde als erster Analyse schritt durchgeführt. Dazu wurden 1-2 µg Prä-tRNA L14 in Nuklease freiem Wasser 5 min bei 55°C Prä-inkubiert und anschließend in 1x F-Puffer (50 mM Tris pH 7, 20 mM NaCl, 20 µg/mL BSA) mit 4.5 mM MgCl<sub>2</sub>, sowie 5 mM DTT und 1 µl gereinigten Enzym bei 37°C inkubiert. Sowohl nach 5 min als auch nach 30 min wurden 8 µl Ansatz entnommen und die Reaktion in 8 µL 2x PPF (eiskalt) abgestoppt. Die Analyse der Spaltung erfolgte durch ein 10%-iges denat. PAA-Gel, Färbung durch GelRed und anschließender UV Beleuchtung.

#### **3.4.8 Aktivitätsanalyse der Proteine (radioaktiv)**

Die Aktivitätsanalyse der Proteine mit radioaktivem Substrat wurde zum einen genutzt, um eine spezifische Spaltung des Substrates durch das Enzym zu beobachten, sowie die Bestimmung der Parameter durch eine Michaelis-Menden Analyse.

Für die radioaktive Untersuchung der Spaltung der Prä-tRNA wurden die am 5'-Ende markierten Substrate (1500 cpm pro Spur) zunächst für 5 min bei 55°C erhitzt und anschließend für die Rückfaltung auf 37°C gebracht. Während der Inkubation des radioaktiven Substrats [S] wurde der Reaktionsansatz für das Enzym vorbereitet. Dabei wurden das Enzym mit Hilfe des Verdünnungspuffers EDB (30 mM Tris-HCl pH 7,8, 30 mM NaCl, 10 µg/mL BSA, 1 mM DTT, 0,3 mM EDTA verdünnt und in einen Reaktionsansatz bestehend aus 1x F-Puffer (50 mM Tris-HCL pH 7, 20 mM NaCl, 4,5 mM MgCl<sub>2</sub>, 20 µg/mL BSA und frisch hinzugegebenes 5 mM DTT) gegeben. Der Ansatz wurde für 5 min bei 37°C prä-inkubiert. Nach Abschluss der Inkubationen wurden 16 µL [E] mit 4 µL [S] vermischt und nach definierten Zeitwerten Proben entnommen. Durch Zugabe der Probe zu eiskaltem Probenpuffer (2x PPF,  $c_{\text{final}} = 1x$ ) wurde die Reaktion abgestoppt. Die so erhaltenen Proben wurden auf einem denaturierendem PAA-

Gel (3.3.6) aufgetrennt und durch das Auflegen einer Imager-plate (Fujifilm) über Nacht entwickelt. Die Imager-plate (Fujifilm) wurde am folgenden Tag durch einen Phosphoimager (Bio-Imaging Analyser BAS 1000, Fujifilm) eingelesen und mit Hilfe der AIDA Software (Aida Image Analyser Version 3.45) ausgewertet.

### 3.4.9 Bestimmung des $K_M$ und $V_{max}$ (*single-turn-over*)

Bei der Bestimmung der Konstanten  $K_M$  und  $V_{max}$ , oder auch  $k_{react}$  genannt unter *single-turn-over* Bedingungen liegt das Protein im deutlichen Überschuss zum Substrat vor ( $[E] \gg [S]$ ). Diese Bedingungen sind unabhängig von der Dissoziation des Produktes.

Für die Experimente wurde ein Enzymsatz [E] in 1x F-Puffer (50 mM Tris-HCl pH 7,0, 20 mM NaCl, 4,5 mM  $MgCl_2$ , 20  $\mu g/mL$  BSA und 5 mM DTT) vorbereitet ( $\frac{4}{5}$  V) und nach einer 5 min Präinkubation mit dem Substratsatz [S] ( $\frac{1}{5}$  V), bestehend aus markierter Prä-tRNA vermischt, welcher für 5 min bei 55°C und 5 min bei 37°C Prä-inkubiert wurde. Die Reaktion wurde nach definierten Zeitpunkten abgestoppt, in dem  $\frac{1}{5}$  V der Reaktion mit demselben Volumen an eiskaltem 2x PPF gemischt wurde. Die Proben wurden über ein 20%-iges denat. PAA-Gel (3.3.6) aufgetrennt und durch das Auflegen einer Imager-plate (Fujifilm) über Nacht entwickelt. Die Imager-plate (Fujifilm) wurde am folgenden Tag durch einen Phosphoimager (Bio-Imaging Analyser BAS 1000, Fujifilm) eingelesen und mit Hilfe der AIDA Software (Aida Image Analyser Version 3.45) ausgewertet. Die Berechnungen erfolgten analog zu [108].

### 3.4.10 Bestimmung des $K_M$ und $V_{max}$ (*multiple-turn-over*)

Bei der Bestimmung der Konstanten  $K_M$  und  $V_{max}$ , oder auch  $k_{react}$  genannt unter *multiple-turn-over* Bedingungen liegt das Substrat im deutlichen Überschuss zum Enzym vor ( $[E] \ll [S]$ ). Dabei wird die Proteinkonzentration konstant gehalten. In diesen Bedingungen spielt die Produktdissoziation eine Rolle.

Für die Experimente wurde ein Enzymsatz [E] (1 nM) in 1x F-Puffer (50 mM Tris-HCl pH 7,0, 20 mM NaCl, 4,5 mM  $MgCl_2$ , 20  $\mu g/mL$  BSA und 5 mM DTT) vorbereitet ( $\frac{4}{5}$  V) und nach einer 5 min Prä-inkubation mit dem Substratsatz [S] ( $\frac{1}{5}$  V), bestehend aus markierter und nicht-markierter Prä-tRNA vermischt,

welcher für 5 min bei 55°C und 5 min bei 37°C Prä-inkubiert wurde. Die Reaktion wurde nach definierten Zeitpunkten abgestoppt, in dem  $\frac{1}{5}$  V der Reaktion mit demselben Volumen an eiskaltem 2x PPF gemischt wurde. Die Proben wurden über ein 20%iges denat. PAA-Gel (3.3.6) aufgetrennt und durch das Auflegen einer Imager-plate (Fujifilm) über Nacht entwickelt. Die Imager-plate (Fujifilm) wurde am folgenden Tag durch einen Phosphoimager (Bio-Imaging Analyser BAS 1000, Fujifilm) eingelesen und mit Hilfe der AIDA Software (Aida Image Analyser Version 3.45) ausgewertet. Die Berechnungen erfolgten analog zu [108].

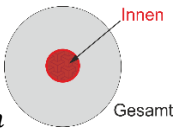
### 3.4.11 RNA-Bindung: DRaCALAs

Das *Differential Radial Capillary Action of Ligand Assay* oder kurz: DRaCALA ist eine schnelle Methode Bindungen zwischen Nukleinsäuren und Proteinen zu untersuchen. Die Methodik beruht auf dem unterschiedlichen Dissoziationsverhalten zwischen den einzelnen Stadien der Bindung beziehungsweise den freien Komponenten der Reaktion [123]. In dieser Arbeit sollte mit dieser Methode die Bindungskonstante zwischen einer Prä-tRNA und AaRP oder den HARPs untersucht werden.

Es wird ein Enzymsatz [E] in EMSA-Puffer (Tabelle 14) vorbereitet und für 5 min bei 37°C Prä-inkubiert. Der Substratsatz wird ebenfalls in EMSA-Puffer vorbereitet und 5 min bei 55°C und 5 min bei 37°C prä-inkubiert. Die Reaktion wird durch das Mischen gleicher Volumina des [E] und [S] gestartet und bei 37°C für 10 min inkubiert. Anschließend wurden 8 µL der Reaktion auf eine Nitrocellulose-Membran getropft und bei RT getrocknet. Die Membran wurde in Folie verpackt und durch auflegen einer Imaging-Platte über Nacht entwickelt. Am nächsten Tag wurde analog zu den Kinetik-Gelen fortgefahren (3.4.9). Die Berechnung der gebundenen Fraktion ( $F_B$ ) wurde mit folgender Formel durchgeführt:

$$F_B = \frac{I_{\text{innen}} - A_{\text{innen}} * \left( \frac{I_{\text{Gesamt}} - I_{\text{innen}}}{A_{\text{Gesamt}} - A_{\text{innen}}} \right)}{I_{\text{Gesamt}}}$$

$I$  = Intensität;  $A$  = Fläche;  $F_B$  = Fraktion, gebunden



### 3.4.12 Oligomerisierungsanalysen mit Glutaraldehyd

Zur Aufklärung des Oligomerisierungszustandes von Proteinen kann eine irreversible Verlinkung mit Glutaraldehyd durchgeführt werden [124]. Die verknüpften Proben können anschließend durch verschiedene Methoden analysiert werden.

Für die Analysen wurden verschiedene Proteinkonzentrationen (5-500 nM) genutzt, sowie diverse Konzentrationen an Glutaraldehyd (0,1 mM – 1 mM). Es wurden 20 µL Protein mit 5 µL Glutaraldehyd gemischt und für eine halbe Stunde bei Raumtemperatur inkubiert. Die Reaktion wurde gestoppt durch die Zugabe von 2 µL 1 M Tris pH 7. Die Proben wurden mit Laemmli Puffer versetzt, für 5 min bei 95°C aufgekocht und anschließend auf ein 12%-iges SDS-Gel aufgetragen. Anschließend wurde ein Western Blot (3.7.8) durchgeführt und mit Hilfe von spezifischen Antikörpern ausgewertet.

## 3.5 Kristallisation von Proteinen

Die Kristallisation von Proteinen dient dazu eine strukturelle Aufklärung durchzuführen. Bei diesem Verfahren werden die Proteine mit unterschiedlichen Puffer-Bedingungen vermischt, so dass sich im Idealfall einige Kristalle bilden, welche sowohl polarisieren als auch im Röntgengerät eine Streuung zeigten. Die Kristallisation wurde in Kooperation mit der AG Reuter (Uni Marburg), sowie der Facility MarXTal (Uni Marburg) durchgeführt.

### 3.5.1 Kristallisation im „*sitting-drop*“ Verfahren

Das *sitting-drop* Verfahren wurde bei der Kristallisation durch MarXTal, sowie bei der AG Reuter angewendet. Dabei wird in eine Kristallisationsplatte (24 well) 500 µL einer Reservoir Lösung geben, sowie 1 µL der Lösung auf den mittleren Steg. Nach dem Vorlegen der Kristallisationspuffer (Tabelle 15) wurde 1 µL des Proteins mit der Lösung vermischt und mit einer Folie abgedeckt. Dies diente zur Verhinderung der Austrocknung des Tropfens. Anschließend wurden die Platten bei 4°C oder 18°C statisch inkubiert. Der Fortschritt der Kristallbildung wurde mit

Hilfe eines Mikroskops, sowie einer integrierten Kamera festgehalten. Die Analyse der Kristalle erfolgte durch Röntgenstrahlung (3.5.3).

*Tabelle 15: Übersicht der Kristallisationsbedingungen*

In dieser Tabelle sind die verschiedenen Kristallisationsbedingungen aufgeführt, welche von MarXtal verwendet worden sind. Angegeben ist zusätzlich das Protein. Alle Ansätze wurden bei 18°C und 4°C durchgeführt

Kristallisations-Kit	Protein	Durgeführt von
<b>AmSO<sub>4</sub></b>	Aq_880/ Hvo_1114	MarXtal
<b>Classics</b>	Aq_880/ Hvo_1114	MarXtal
<b>Classics Lite</b>	Aq_880/ Hvo_1114	MarXtal
<b>Core I</b>	Aq_880/ Hvo_1114	MarXtal
<b>Core II</b>	Aq_880/ Hvo_1114	MarXtal
<b>Core III</b>	Aq_880/ Hvo_1114	MarXtal
<b>Core IV</b>	Aq_880/ Hvo_1114	MarXtal
<b>JCSG+</b>	Aq_880/ Hvo_1114	MarXtal
<b>MBC II</b>	Aq_880/ Hvo_1114	MarXtal
<b>Morpheus</b>	Aq_880/ Hvo_1114	MarXtal
<b>Morpheus II</b>	Aq_880/ Hvo_1114	MarXtal
<b>PACT</b>	Aq_880/ Hvo_1114	MarXtal
<b>Anions</b>	Hvo_1114	MarXtal
<b>Cryos</b>	Hvo_1114	MarXtal
<b>JB Penta</b>	Hvo_1114	MarXtal
<b>MBC I</b>	Hvo_1114	MarXtal

### 3.5.2 Kristallisation im „*hanging-drop*“ Verfahren

Bei der Kristallisation im *hanging-drop* Verfahren wird die Kristallisation auf einem Deckglas durchgeführt, welches über der Reservoir Lösung hängt. Diese Methode wurde gewählt, nachdem im *sitting-drop* Verfahren erste Kristalle detektiert wurden. Vorteil bei dieser Methode ist, dass der Kristall nicht an der Oberfläche der Platte oder des Deckglases adhären kann.

Das Experiment wurde analog zu der *sitting-drop* Methode durchgeführt (3.5.1), allerdings wurde keine Folie zum Verschließen der Platte verwendet, sondern das mit Fett versehene Deckglas.

### 3.5.3 Analyse der Kristalle im Röntgengerät

Die Analyse der Kristalle erfolgte zunächst im Fachbereich am Hauseigenen Röntgengerät. Die Kristalle wurden mit Hilfe von kleinen Nylonschlaufen in verschiedenen Größen (Hampton Research) aufgenommen und direkt in flüssigen Stickstoff eingefroren. Die Datensammlung erfolgte mit einem MAR-Incoatec MAR345 Image Plate Detektor mit I $\mu$ <sup>TM</sup> Mikrofokusröhre und Oxford Cryosystems Tieftemperaturkühlung. Die Sammlung erfolgte zunächst nur bei zwei Winkeln. Nach der Auswertung der Bilder mit der dazugehörigen Software wurden die Kristalle entweder verworfen oder weiter in flüssigem Stickstoff gelagert.

## 3.6 Chromatographie

### 3.6.1 Affinitätschromatographie

Die Affinitätschromatographie basiert auf der Affinität des Liganden entweder zu einem Antikörper oder eines Cofaktors. In den meisten Fällen ist der Antikörper bzw. Cofaktor an eine Agarosematrix gebunden und der Ligand wird in Lösung dazu geben, meist aus einem Zellextrakt.

In dieser Arbeit wird die Affinität von einer Proteinfole (Histidin) zu Nickel-Ionen ausgenutzt. Dabei erfolgt das Lösen des Proteins von den Nickel-Ionen durch eine Erhöhung der Imidazol Konzentration im Puffer. Für die Reinigung der Proteine wurde ein Natrium-Phosphat-Puffer mit Imidazol (NPI, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20-500 mM Imidazol, pH 8) verwendet. Bei einer denaturierenden Reinigung wurden dem Puffer 8 M Urea beigesetzt. Die Affinitätsreinigung mittels einer FPLC wurde sowohl für analytische als auch präparative Reinigungen des entsprechenden Proteins genutzt. Es wurden fertig gepackte Säulen (HisTrap<sup>TM</sup>HP 1 mL, GE Healthcare) genutzt.

Die (aufgetauten) Zellen wurden in 20 - 50 mL NPI-20 aufgenommen und mit Hilfe von Ultraschall aufgeschlossen (3x; Intensität von 50-60%, Output 5), wobei die Zellen zwischen den Zyklen auf Eis aufbewahrt wurden. Der

Zellaufschluss erfolgte eisgekühlt, mit Ausnahme des Aufschlusses von Expressionskulturen von Aq\_880. Anschließend erfolgte die Abtrennung der unlöslichen Bestandteile (4°C, 8 000 rpm, 30 min). Vor dem Laden der Säule wurde die Probe durch einen 0.22 µm Filter gegeben, um mögliche Rückstände zu entfernen.

Die Säule wurde bei einem Fluss von 1 mL/min mit 5 Column Volumes (CV) NPI-20 äquilibriert. Bei größeren Volumina an Zellaufschluss wurde das Beladen der Säule mit Hilfe eines Perfusor (Braun) bei einem Fluss von 60 mL/h durchgeführt. Bei kleineren Volumina wurde ein Superloop (10 mL) verwendet, welcher mit der FPLC gesteuert werden konnte. Anschließend wurde die Säule bis zum Erreichen einer stabilen Basislinie mit NPI-20 gewaschen. Die Elution erfolgte über 30 CV mittels eines linearen Gradienten bis 500 mM Imidazol bei einer Flussrate von maximal 2 mL/min. Nach der Etablierung der Reinigung wurde gegebenenfalls ein Stufengradient genutzt, um eine höhere Reinheit zu erreichen. Die Elution erfolgte in 1-3 mL Fraktionen, welche anschließend mittels einer SDS-PAGE (3.7.6) und Coomassie Färbung (3.7.7), sowie später durch das Stain-Free System (BioRad) analysiert wurden. Die Säule wurde nach Herstellerangaben gereinigt und gelagert.

### **3.6.2 Anionenaustauschchromatographie**

Die Ionenaustauschchromatographie ist eine Trennmethode welche auf der Ladung der Proteine basiert. Vorteil dieser Methode ist, dass ein natives Protein, ohne Fusionen, verwendet werden kann. Bei dieser Methode bilden sich reversible heteropolare Bildungen zwischen den Matrix-Ionen (Fest) und des Protein-Lysats (Gegen-Ionen, Flüssig). Die Elution erfolgt bei dieser Methode durch eine schrittweise Erhöhung der Gegen-Ionen.

In dieser Arbeit wurden folgende Säulen verwendet: MonoQ™ 5/50 GL (starker Anionenaustauscher, GE Healthcare) und DEAE-Sepharose® (schwacher Anionenaustauscher, Sigma-Aldrich). Für die Chromatographie wurde ein 20 mM Tris Puffer pH 8 mit 0-1 M NaCl verwendet.

Bei der Verwendung der MonoQ™ 5/50 GL Säule (1 mL CV) lag bereits ein zuvor durch Affinitätsreinigung behandeltes Protein Lysat vor, welches über Nacht in den MonoQ-Start-Puffer dialysiert (Tabelle 14) wurde. Das Protein-Lysat

wurde durch einen 0,22 µm Filter gegeben, um mögliches Präzipitat abzutrennen. Die Säule wurde mit 5 CV MonoQ-Start-Puffer equilibriert (2 mL/min), anschließend wurde die Probe mit 1 mL/min geladen und bis zum Erreichen einer stabilen Basislinie gewaschen (2 mL/min). Die Elution des Proteins erfolgte durch eine Steigerung des NaCl von 0 auf 1 M über 20 CV, wobei die Proben in 0.5 - 1 mL Fraktionen gesammelt wurden. Die Fraktionen wurden über ein SDS-Gel (3.7.6) mit anschließender Coomassie-Färbung (3.7.7) analysiert. Ausgewählte Proben wurden über Nacht in den entsprechenden Puffer dialysiert (3.4.3). Die Reinigung der Säule erfolgte nach Herstellerangaben.

Bei Verwendung der DEAE-Sepharose® wurde eine selbstgepackte Säule von 20 mL CV verwendet. Diese Reinigungsmethode war der erste Schritt einer mehrstufigen Reinigung von rekombinantem nativem Aq\_880 (ohne Fusion). Dabei wurden 1 L Kulturvolumen angezogen (3.4.1) und das Pellet in 20 mM Tris pH 8 Puffer aufgenommen, die Zellen aufgeschlossen (3.4.2) und das Protein-Lysat geklärt (3.4.2). Die Säule wurde für 3 CV mit 20 mM Tris pH 8 equilibriert und die Probe geladen (3 mL/min). Anschließend erfolgte die Elution der Proteine durch einen linearen Gradienten gegen 1 M NaCl über 10 CV. Es wurden 3 mL Fraktionen gesammelt, über ein SDS-Gel (3.7.6) mit Coomassie-Färbung (3.7.7) analysiert, sowie später durch das Stain-Free System (BioRad) und anschließend in den entsprechenden Lagerpuffer (Tabelle 14) dialysiert (3.4.3).

### 3.6.3 Kationenaustauschchromatographie

Die Kationenaustauschchromatographie wurde verwendet, um die T7 RNA Polymerase zu reinigen. Dabei wurde ein bereits vorgereinigtes Protein-Lysat verwendet und durch einen 0,22 µm Filter gegeben. Die Fractogel® EMD SO<sub>3</sub><sup>-</sup> (Merck)- Säule wurde mit C-100 Puffer (20 mM Na-Phosphat pH 7,7; 1 mM Na<sub>2</sub>EDTA, 1 mM DTT; 5% Glycerin (v/v); 2 µg/mL PMSF, 100 mM NaCl) equilibriert (2 mL/min), die Probe mit 1 mL/min geladen und anschließend bis zum Erreichen einer stabilen Basislinie mit C-100-Puffer gewaschen (2 mL/min). Die Elution erfolgte als linearer Gradient, wobei die NaCl Konzentration bis auf 1 M erhöht wurde (C-Puffer mit 1 M NaCl). Die Fraktionen wurden in 3 mL



gesammelt und anschließend durch ein SDS-Gel (3.7.6) und eine Coomassie-Färbung (3.7.7), sowie später durch das Stain-Free System (BioRad) analysiert. Ausgewählte Fraktionen wurden zusammen gegen den T7-RNAP-Lagerpuffer (Tabelle 14) dialysiert (3.4.3).

### 3.6.4 Größenausschlusschromatographie

Die Trennung von Proteinen mittels Größenausschlusschromatographie basiert auf der unterschiedlichen Größe von Proteinen. Dabei werden die Proteine durch eine poröse Gelmatrix gegeben, wobei kleinere Proteine stärker und große Proteine schwächer mit dieser interagieren. Durch die unterschiedlich starke Interaktion mit der Matrix wird die Elution verzögert. Große Proteine gelangen schneller, kleine langsamer hindurch. Anhand einer Kalibrierkurve mit bekannten Proteinen, kann durch das Elutionsvolumen auf die Größe des unbekannten Proteins geschlossen werden.

In dieser Arbeit wurde die Größenausschlusschromatographie angewendet, um den Oligomerisierungszustand der entsprechenden Proteine unter verschiedenen Bedingungen zu untersuchen. Dazu wurde eine Superose 6 10/300 GL (GE Healthcare) mit 24 mL CV verwendet und einer Trennleistung von 5 000 bis 5 000 000, sowie einem Ladungsvolumen von 250  $\mu$ L.

Die Säule wurde mit 2 CV Storage Buffer (ohne Glycerin; Tabelle 14) equilibriert (maximal 0,4 mL/min). Anschließend wurden die Standardproteine (Apo ferritin Horse spleen 25 mg/mL; Kit for Molecular weights 12 – 200 kDa (Sigma-Aldrich) in drei Mischungen von zwei Proteinen (je ~0.7 mg) eingewogen und in 300  $\mu$ L Storage Buffer gelöst. Die Proteingemische wurden durch einen 0.22  $\mu$ m Filter (Ultrafree®-MC Filter Devices, 0.22  $\mu$ m, Merck) gegeben (3 min, 10 000 rpm) um mögliche Präzipitate, sowie unlösliches Protein abzutrennen.

Das Laden der Standardproteine erfolgte mit einem Volumen von 250  $\mu$ L nach dem Erreichen einer stabilen Basislinie (280 nM) (maximal 0,4 mL/min). Die Elution erfolge über 1 CV mit Storage Buffer (maximal 0,4 mL/min). Nach der Elution wurde erneut bis zum Erreichen einer stabilen Basislinie (280 nM) gewaschen bevor eine neue Probe geladen wurde. Nach Abschluss der Standardproteine wurde über die Software Unicorn 5.01 (GE Healthcare) mit Hilfe der Funktion  $\int$  Peak Integrate das Elutionsvolumen bestimmt. Für die

Berechnung der Kalibrierkurve wurden die molare Masse logarithmisch gegen das Elutionsvolumen aufgetragen und eine exponentielle Trendlinie berechnet (Excel, Microsoft).

Das Laden der zu untersuchenden Proteine erfolgte wie oben für die Standardproteine beschrieben, jedoch wurden hier zwischen 0,6 mg und 10 mg für die Charakterisierung genutzt. Dabei wurden die Proteine ebenfalls in Storage Buffer (ohne Glycerin) verdünnt. Anschließend wurde ebenfalls das Elutionsvolumen mit Hilfe des Programms bestimmt und unbekannte  $x$  in die Formel eingesetzt. Das Ergebnis ( $y$ ) dieser Gleichung entspricht der Molaren Masse des Proteins.

Die Säule wurde zum Schluss nach Herstellerangaben gereinigt, wobei nach 20 Läufen zusätzlich ein Pepsin-Verdau, sowie ein Filterwechsel vorgenommen wurde.

*Tabelle 16: Auflistung der Standardproteine für die Größenausschlusschromatographie*

Die Standardproteine wurden von Sigma-Aldrich bezogen und mit ~0.7 mg für die Größenausschlusschromatographie eingesetzt. Zwei Proteine (groß und klein) sind in einem Mix über die Säule gegeben worden.

Standardmix	Standardprotein	Molare Masse
Mix A	Apoferritin	443 kDa
	BSA	66 kDa
Mix B	$\beta$ -Amylase	200 kDa
	Carbon-Anhydrase	29 kDa
Mix C	Alkoholdehydrogenase	150 kDa
	Cytochrom C	12 kDa

### 3.7 *In vivo* Charakterisierung von Proteinen

Neben den *in vitro* Charakterisierungen der unterschiedlichen Proteine (3.4) wurden diese auch *in vivo* untersucht. Dazu wurde sowohl ein Modellsystem für Gram-negative (3.7.1), als auch Gram-positive Bakterien (3.7.2) verwendet. Die Analyse der Protein-basierten RNase P Varianten wurde in Gram-negative

Bakterien und der P Proteine der RNase P in Gram-positiven Bakterien durchgeführt.

### 3.7.1 Komplementationsstudien mittels *E. coli* BW

Für die Komplementationsstudien wurde ein zuvor erstellter und bereits publizierter *E. coli* Stamm verwendet [113]. Bei diesem Stamm steht die genomische P RNA unter der Kontrolle eines Arabinose-Promotors. Zusätzlich besitzt dieser Stamm einen Selektionsmarker für Chloramphenicol. Wird dieser Stamm mit der Kohlenstoffquelle Glukose, anstelle von Arabinose angezogen, ist er nicht lebensfähig. Um die Funktionalität verschiedener P RNAs, sowie Protein-basierten RNase Ps zu testen, wurden diese auf einem Vektor unter Kontrolle eines Spac-Promotors angeboten. Bei einer funktionalen P RNA, beziehungsweise einer Protein-basierten RNase P ist der Stamm weiterhin, auch unter Glukose Lebensfähig.

Für die statischen Analysen wurden die entsprechenden Stämme auf Selektionsplatten, welche Arabinose enthielten (3.1.1) ausgestrichen und über Nacht bei 37°C inkubiert. Am folgenden Tag wurde eine Einzelkolonie in 500-800 µL LB Medium (ohne Zusätze) aufgenommen, bei 6 000 rpm für 2 min gewaschen und erneut in 800 - 1000 µL LB Medium resuspendiert. Es wurden dreimal 10 µL der Suspension auf getrocknete Selektionsplatten (mind. 30 min), entweder mit Arabinose (Viabilitätskontrolle) oder Glukose (Funktionalitätskontrolle) getropft, für 30 min statisch bei Raumtemperatur und anschließend für mehrere Tage bei 37°C weiter inkubiert. Die Platten wurden mit Hilfe einer Geldokumentationsanlage dokumentiert, dabei wurden die Einstellungen der Blende, Schärfe, Vergrößerung, sowie Belichtungszeit für eine Experimenten Reihe konstant gehalten.

Die nicht-statischen Versuche wurden mit Hilfe des Multiskan™ GO (Thermo Scientific) und der entsprechenden Software (SkanIt™) in einer 96 well Platte durchgeführt. Dabei wurde mit den Proben verfahren, wie bereits oben beschrieben. Allerdings wurden 50 µL der Suspension mit 50 µl 2x Komplementation-Medium (LB mit amp 200, cam 68, ara 20 bzw. glu 20) pro well versetzt. Die Zelldichte wurde automatisiert bei 600 nM alle 15 min bestimmt, wobei zwischen den Messungen die Platte durch Schütteln und bei 37°C

inkubiert wurde. Die Auswertung der Ergebnisse erfolgte in Excel (Microsoft), wobei die Zelldichte gegen die Zeit aufgetragen wurde. Anhand des entstandenen Graphen können Wachstumsunterschiede der einzelnen Stämme miteinander verglichen werden.

### 3.7.2 Komplementationsstudien mittels *B. subtilis* d7

Die Komplementationsstudien mit *B. subtilis* d7 wurden analog zu den Experimenten mit *E. coli* BW im statischen System durchgeführt (3.7.1). Allerdings erfolgte die Transformation der entsprechenden Vektoren mit Hilfe der Elektroporation (3.1.5).

### 3.7.3 Analyse der Protein-Expression in *E. coli* BW

Die Analyse der Proteinexpression wurde durchgeführt, um nachzuweisen, dass die Proteine (HARPs und AaRP) in einer löslichen Form in *E. coli* BW vorlagen. Es wurden 50 mL einer Übernachtskultur unter Arabinose angelegt und bei 37 °C bei 220 rpm inkubiert. Am nächsten Tag wurden die Zellen bei 6 000 rpm und 4°C geerntet, bevor die Zellen in 800 µL Lysepuffer (25 mM Tris-HCl pH 8, 100 mM NaCl, 2 mM EDTA, 0,1 %Triton X-100, 1 Tablette Complete Mini (Roche)) resuspendiert und mit 36 µL Lysozym (100 mg/mL) für 15 min auf Eis inkubiert. Anschließend wurde die Zellsuspension in 1,5 mL Röhrchen mit *Lysing-Matrix B* (MP) überführt und mit Hilfe der FastPrep (MP) dreimal für 20 s aufgeschlossen. Um die Lysematrix und das Lysat zu trennen wurden die Probe für 5 min bei 1 000rpm zentrifugiert, der Überstand in ein neues Reaktionsgefäß überführt und zur Trennung der löslichen und unlöslichen Proteine für 30 min bei 10 000 rpm und 4 °C zentrifugiert.

Die löslichen Proteine befinden sich im Überstand und werden in ein neues Reaktionsgefäß überführt, das Volumen bestimmt und die unlöslichen Proteine im gleichen Volumen an TUS-Puffer (50 mM Tris-HCl pH 7,6, 100 mM NaCl, 8 M Urea) resuspendiert. Die Konzentration der löslichen Gesamtproteine wurde mit Hilfe eines Bradfords (3.4.6) bestimmt und insgesamt 30 µg auf ein TGX-Stain-Free Gel (3.7.9) geladen. Für die unlöslichen Proteine wurde das

gleiche Volumen aufgetragen. Nach der Dokumentation des Gesamtproteins (3.7.9) wurde das Gel geblottet (3.7.8) und ausgewertet.

#### **3.7.4 Analyse der Protein-Expression in *B. subtilis* d7**

Die Analyse der Proteinexpression wurde durchgeführt, um nachzuweisen, dass die RNase P Proteine in einer löslichen Form in *B. subtilis* d7 vorlagen. Für die Analyse wurden 50 mL LB Kultur mit 2% Xylose mit einer Einzelkolonie beimpft, über Nacht bei 37 °C und 200 rpm inkubiert. Am nächsten Tag wurden die Zellen bei 6 000 rpm und 4 °C geerntet, bevor die Zellen in 800 µL Lysepuffer (25 mM Tris-HCl pH 8, 100 mM NaCl, 2 mM EDTA, 0,1 % Triton X-100, 1 Tablette Complete Mini (Roche)) resuspendiert und mit 36 µL Lysozym (100 mg/mL) für 15 min auf Eis inkubiert. Anschließend wurde die Zellsuspension in 1,5 mL Röhrchen mit *Lysing-Matrix B* (MP) überführt und mit Hilfe der FastPrep (MP) dreimal für 20 s aufgeschlossen. Um die Lysematrix und das Lysat zu trennen wurden die Probe für 5 min bei 1 000 rpm zentrifugiert, der Überstand in ein neues Reaktionsgefäß überführt und zur Trennung der löslichen und unlöslichen Proteine für 30 min bei 10 000 rpm und 4 °C zentrifugiert.

#### **3.7.5 Analyse von Fusionsproteinen in Komplementationsstudien**

Die Proben werden zuvor wie unter 3.7.3 und 3.7.4 behandelt. Die löslichen Proteine befinden sich im Überstand welches mit Protino® Ni-NTA Matrix (50 µL) versetzt wurde. Die Proteinsuspension wurde über Kopf rotiert bei 4°C. Anschließend wurde die Matrix abzentrifugiert (2000 rpm, 4°C) und der Überstand verworfen. Die Matrix wurde zweimal mit Lysepuffer gewaschen, bevor die Proteine mit 50 µL NPI-500 eluiert wurden. Sowohl von den löslichen Proteinen, als auch von den unlöslichen Proteinen (gelöst in TUS-Puffer) wurden insgesamt 10 µL auf ein 12% TGX-Stain Free Gel aufgetragen, mit der ChemiDoc (BioRad) dokumentiert (3.7.9) und anschließend geblottet und ausgewertet (3.7.8).

### 3.7.6 SDS-PAGE

Die SDS-PAGE wurde durchgeführt, um Proteine ihrer Größe nach aufzutrennen. Dabei wurden 10-15%-ige Gele verwendet, sowie das Gelgießsystem von BioRad mit Mini-Gelen und einer Gel-Dicke von 0,75 mm.

Zuerst wurde das Trenngel vorbereitet und gegossen, dabei wurden 1x Trenngelpuffer (0,375 M Tris-HCl pH 8,8, 0,15% SDS) mit Acrylamid/Bisacrylamid für die entsprechende Dichte gemischt und durch die Zugabe von APS (10%) und TEMED auspolymerisiert. Das Trenngel wurde während der Polymerisierung mit Isopropanol überschichtet. Anschließend wurde das Sammelgel gegossen, dabei wurden 1x Sammelgelpuffer (0,125 M Tris-HCl pH 6,8, 0,15% SDS) mit Acrylamid/Bisacrylamid ( $C_{\text{final}}$  4%) gemischt und die Polymerisation ebenfalls mit APS (10%) und TEMED gestartet. Es wurde vor dem finalen auspolymerisieren noch ein Kamm eingesteckt (10 oder 15 Taschen). Die Lagerung der Gele erfolgte mit etwas Feuchtigkeit bei 4°C für maximal zwei Wochen.

Das Auftrennen der Proben erfolgte zunächst durch das versetzen der Proteine mit 1x Laemmli-Puffer (4x Puffer: 100 mM Tris-HCl pH 6,8, 8% (w/v) SDS, 8% (v/v)  $\beta$ -Mercaptoethanol, 0,04% (w/v) Bromphenolblau, 30% (v/v) Glycerin) und optionalen aufkochen bei 95 - 98°C für 3 - 5 min. Der Lauf wurde in 1x SDS-Laufpuffer (10x SDS-Laufpuffer (Roth)) bei maximal 180 V durchgeführt, bis die Lauffront das Gel-Ende erreichte. Die Visualisierung der Proben erfolgte entweder durch eine Coomassie-Färbung (3.7.7) oder einem Western Blot (3.7.8).

### 3.7.7 Coomassie-Färbung

Die Coomassie Färbung dient zum visualisieren aufgetrennter Proteine durch eine SDS-Page (oder TGX). Dabei wurde das Gel für 30 - 60 min in Coomassie-Lösung (0,2% (w/v) Coomassie Brilliant Blue G-250, 10% (v/v) Essigsäure, 50% (v/v) Methanol) inkubiert und anschließend entfärbt. Dieser Prozess wurde entweder durch mehrmaliges Aufkochen in Wasser oder dem langsamen entfärben in Wasser über Nacht durchgeführt. Anschließend erfolgte die fotografische Dokumentation.

### 3.7.8 Western Blot

Der Western Blot ist eine Methode, um eine geringe Proteinkonzentration mit spezifischen Antikörpern zu visualisieren. Dazu wird zunächst ein SDS-Gel genutzt, um die Proteine aufzutrennen. Anschließend erfolgt der Transfer der Proteine auf eine PVDF-Membran (0,45 µM, Immobilon-P (Merck)) mit Hilfe des Trans-Blot® SD Semi-Dry Transfer Cell (BioRad) oder Trans-Blot Turbo Transfer System (BioRad). Dabei wurde für den Trans-Blot eine Spannung von 12 V für 30 - 40 min angelegt. Im Turbo-System wurden 25 V für 3 - 10 min genutzt. Der Blot wurde wie folgt aufgebaut: zwei in Transferpuffer getränkte Whatman-Papiere (48 mM Tris, 39 mM Glycin, 20% (v/v) Methanol, 0,04% (v/v) SDS) wurden um die Membran mit dem in Transferpuffer äquilibrierten SDS-Gel geschichtet. Die Membran wurde zuvor nach Herstellerangaben aktiviert. Der Blot mit einen TGX-Stain-Free Gelen erfolgte ähnlich zu SDS-Gelen. Allerdings wurde das Gel nicht äquilibriert, sondern direkt eingesetzt. Der Blot-Vorgang wurde für 3-10 min bei 25 V durchgeführt.

### 3.7.9 TGX Stain-Free Gele

Das Auftrennen von Proteinen mit TGX-Gelen kann zusätzlich zu SDS-Gelen genutzt werden. Die TGX-gele von BioRad können mit einer höheren Spannung gefahren werden und wenn es Stain-Free Varianten sind, können die Proteine direkt in der ChemiDoc visualisiert werden. Es bedarf also keiner Färbung der Gele. Das Visualisieren im Stain-Free Modus wird durch die Aktivierung von aromatischen Aminosäuren ausgelöst.

Es wurden Gele mit 10% oder 12% genutzt und nach Herstellerangaben gegossen. Der Lauf erfolgte analog zu SDS-Gelen (3.7.6), allerdings mit einer Spannung von 250 - 300 V und einer Laufzeit von 15 - 30 min.

## 4 Ergebnisse und Diskussion

In den nachfolgenden Kapiteln finden sich die Zusammenfassungen bereits veröffentlichter Publikationen (Publikation 1 und 2), einem Artikel in einer Enzyklopädie (Publikation 3) sowie bisher nicht publizierter Ergebnisse

Jede Zusammenfassung beinhaltet jeweils den Verweis auf einzelnen Abbildungen und Tabellen in den entsprechenden Publikationen. Zusätzlich ist die Seitenzahl angegeben, wo die jeweiligen Abbildungen und Tabellen der Publikation zu finden sind. Die verwendeten Methoden und Referenzen sind in den jeweiligen Manuskripten aufgeführt. Detaillierte Angaben zu den Experimenten befinden sich teilweise auch im Material und Methoden Teil dieser Arbeit.

### 4.1 Publikation 1:

#### **Minimal and RNA-free RNase P in *Aquifex aeolicus***

In dem hyperthermophilen Organismus *A. aeolicus* konnten auch nach intensiver bioinformatischer Suche keine Gene für die RNase P Untereinheiten identifiziert werden. In Bakterien wird dieses essenzielle Enzym durch eine RNA-Untereinheit (*rnpB*) und eine Proteinuntereinheit (*rnpA*) aufgebaut. Es wurde spekuliert, dass eine neuartige Form der RNase P in diesem Organismus vorliegen könnte, da in früheren Experimenten eine Prozessierung der 5'-Flanke nachgewiesen wurde.

Zunächst wurden durch mehrstufige chromatographische Aufreinigungen Fraktionen identifiziert, welche eine RNase P Aktivität aufwiesen (Fig. 1 (Seite 71); Fig. S1-S10 (Seiten 78-85)). Diese wurden zuerst auf die Proteinzusammensetzungen untersucht, wobei drei Proteine als mögliche RNase P-Kandidaten in Frage gekommen sind: Aq\_221, Aq\_707 und Aq\_880. Es zeigte sich, dass nur Aq\_880 in der Lage war, die 5'-Prozessierung von Prä-tRNAs beeinflussen (Fig. 2 (A) (Seite 72); Fig. S13 (Seite 88)). Überraschenderweise war dieses kleine Protein (~23 kDa) allein dazu in der Lage die Funktion einer RNase P zu übernehmen (Fig. 2 (A) (Seite 72)). Mehrere Kontrollen (Fig. 2 (F) (Seite 72); Fig. S11-S12, S14 (Seiten 86,87,89) verifizierten dass es sich bei der RNase P aus *A. aeolicus* um ein reines Proteinenzym ohne



RNA-Beteiligung handelt. Diese Entdeckung stellte ein Novum dar, denn es wurde bis zu diesem Zeitpunkt angenommen, dass Protein-basierte Form der RNase P exklusiv in Eukarya zu finden sind. Die Untersuchung des aktiven Zentrums von Aq\_880 zeigte, dass die Organisation der katalytisch wichtigen Aspart-Reste ähnlich zu der im bereits charakterisierten singulären protein-basierten RNase P-Enzym aus *Arabidopsis thaliana* (PRORP) ist (Fig. S17-S20 (Seiten 92-94)). Zu den *in vitro* Untersuchungen konnte zusätzlich gezeigt werden, dass dieses Enzym in der Lage ist, die RNase P von *E. coli*, sowie auch die komplexere Form in Hefe *in vivo* funktional zu ersetzen (Seite 73), allerdings mit einem Defizit im Wachstum.

Mittels enzymkinetischer Untersuchungen wurden die kinetischen Parameter  $K_M$  und  $k_{\text{react}}$  bestimmt, welche mit denen von bereits bekannten RNase P-Enzymen vergleichbar sind (Fig. 2 (D-E) (Seite 72)). Durch bioinformatische Datenanalyse konnten auch in anderen Prokaryoten Homologe von Aq\_880 identifiziert werden (Fig. 4 (Seite 74); Fig. S23 (Seiten 97-99)). Dabei zeigte sich, dass alle identifizierten Organismen zudem noch eine RNA-basierte RNase P besitzen, so dass hier zwei potenzielle RNase P-Enzymvarianten in einer Zelle vorliegen (Table S1 (Seiten 100-103)). Erste enzymkinetische Analysen zeigten, dass auch die Aq\_880-homologen Proteine aus dem Bakterium *Thermodesulfatator indicus* und dem Archaeon *Methanothermobacter thermautotrophicus*, in der Lage sind die Spaltung der 5'-Flanke einer Prä-tRNA zu katalysieren. Zusätzlich konnte gezeigt werden, dass die klassische RNA-basierte Form von *T. indicus* katalytisch ebenfalls aktiv ist. Die RNase P von *M. thermautotrophicus* wurde bereits in früheren Publikationen bezüglich ihrer katalytischen Aktivität identifiziert und charakterisiert (Table 1 (Seite 74); Fig. S21 (Seite 96)).

#### 4.1.1 Originalpublikation

Auf den nächsten Seiten befindet sich die Originalpublikation „*Minimal and RNA-free RNase P in Aquifex aeolicus*“, welche im September 2017 in *Proceedings of the National Academy of Sciences of the United States of America* (PNAS) veröffentlicht worden ist.



# Minimal and RNA-free RNase P in *Aquifex aeolicus*

Astrid I. Nickel<sup>a</sup>, Nadine B. Wäber<sup>a</sup>, Markus Göbringer<sup>a</sup>, Marcus Lechner<sup>a</sup>, Uwe Linne<sup>b</sup>, Ursula Toth<sup>c</sup>, Walter Rossmann<sup>c</sup>, and Roland K. Hartmann<sup>a,1</sup>

<sup>a</sup>Institute of Pharmaceutical Chemistry, Philipps-Universität Marburg, 35037 Marburg, Germany; <sup>b</sup>Faculty of Chemistry, Mass Spectrometry, Philipps-Universität Marburg, 35032 Marburg, Germany; and <sup>c</sup>Center for Anatomy & Cell Biology, Medical University of Vienna, 1090 Vienna, Austria

Edited by Sidney Altman, Yale University, New Haven, CT, and approved September 11, 2017 (received for review May 12, 2017)

**RNase P is an essential tRNA-processing enzyme in all domains of life. We identified an unknown type of protein-only RNase P in the hyperthermophilic bacterium *Aquifex aeolicus*: Without an RNA subunit and the smallest of its kind, the 23-kDa polypeptide comprises a metallo-nuclease domain only. The protein has RNase P activity in vitro and rescued the growth of *Escherichia coli* and *Saccharomyces cerevisiae* strains with inactivations of their more complex and larger endogenous ribonucleoprotein RNase P. Homologs of *Aquifex* RNase P (HARP) were identified in many Archaea and some Bacteria, of which all Archaea and most Bacteria also encode an RNA-based RNase P; activity of both RNase P forms from the same bacterium or archaeon could be verified in two selected cases. Bioinformatic analyses suggest that *A. aeolicus* and related *Aquificaceae* likely acquired HARP by horizontal gene transfer from an archaeon.**

protein-only RNase P | *Aquifex aeolicus* | tRNA processing | HARP

The architectural diversity of RNase P enzymes is unique: In Bacteria, Archaea, and in the nuclei and organelles of many Eukarya, RNase P is a complex consisting of a catalytic RNA subunit and a varying number of proteins (one in Bacteria, at least four in Archaea, and up to 10 in Eukarya) (1, 2). A different type of RNase P was discovered more recently in human mitochondria (3) and, subsequently, in land plants and some protists (4, 5). This form, termed proteinaceous or protein-only RNase P (PRORP), lacks any RNA subunit and consists of one or three (animal mitochondria) protein subunit(s); it is found in most branches of the eukaryotic phylogenetic tree (6).

Bacterial RNase P enzymes identified so far are composed of a ~400-nt-long catalytic RNA subunit (encoded by *mpB*) and a small protein subunit of ~14 kDa (encoded by *mpA*) (7). However, no *mpA* and *mpB* genes were identified in the genome of *Aquifex aeolicus* or other *Aquificaceae* (8–12). The genetic organization of *A. aeolicus* tRNAs in tandem clusters and as part of ribosomal operons and the detection of tRNAs with canonical mature 5'-ends in total RNA extracts from *A. aeolicus* implied the existence of a tRNA 5'-maturation activity (9) that was indeed subsequently detected in cell lysates of *A. aeolicus* (11, 13). However, to date, the identity and biochemical composition of RNase P in *A. aeolicus* has remained enigmatic.

## Results and Discussion

Here, we pursued a classical biochemical approach to identify the RNase P of *A. aeolicus*. The purification procedure consisted of three consecutive chromatographic steps: anion exchange, hydrophobic interaction, and size exclusion chromatography (AEC, HIC, and SEC, respectively; Fig. 1A and SI Appendix, Figs. S1–S8). RNase P activity was assayed at all purification steps. To identify putative protein components of the enzyme, fractions with low and high RNase P activity from different purification steps were comparatively analyzed by step-gradient SDS/PAGE, and protein bands correlating with activity (Fig. 1B) were subjected to mass spectrometry. An example is SEC (SI Appendix, Fig. S5) fraction B4 displaying maximum activity combined with an enrichment of protein bands 1 and 2 identified as the hypothetical protein Aq\_880 and polynucleotide phosphorylase (PNPase), respectively

(Fig. 1B). Several protein bands correlating less strongly with RNase P activity were identified as well. The most abundant proteins in fractions containing highest RNase P activity (SI Appendix, Figs. S9 and S10), as inferred from mass spectrometry, were in the order of decreasing peptide representation: (i) Aq\_880, (ii) glutamine synthetase, (iii) ribosomal protein S2, (iv) PNPase, (v) N utilization substance protein B homolog (NusB), and (vi) Aq\_707 (with similarity to an *Escherichia coli* tRNA binding protein of the MnmC family). Furthermore, Aq\_880 was the only protein that was also found in an HIC fraction with low RNase P activity eluting at 0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (SI Appendix, Figs. S3 and S10). The presence of ribosomal protein S2 combined with a previously described association of RNase P with 30S ribosomal subunits in *Bacillus subtilis* (14) suggested a possible association of *A. aeolicus* RNase P with the ribosome, which, however, could not be confirmed experimentally.

The 22.6-kDa protein Aq\_880 was recombinantly expressed in *E. coli* and affinity-purified using a C-terminal His tag. Addition of affinity-purified Aq\_880 protein to an active *A. aeolicus* HIC fraction not only boosted RNase P activity, but recombinant Aq\_880 alone was able to process the precursor tRNAs specifically and efficiently at the canonical RNase P cleavage site without requiring any additional components (Fig. 2A). We further analyzed whether processing by recombinant Aq\_880 generates tRNA with a 5'-phosphate end, as any other previously characterized RNase P. TLC indeed confirmed a 5'-phosphate at the 5'-terminal G<sub>+1</sub> residue of tRNA<sup>Gly</sup> processed by recombinant Aq\_880 (Fig. 2B and C). A contamination of the recombinant Aq\_880 preparation with endogenous *E. coli* RNase P was excluded by RT-PCR analysis and micrococcal nuclease pretreatment (SI Appendix, Figs. S11 and S12). Since Aq\_707 and PNPase (Aq\_221) copurified with RNase P activity and peptides matching Aq\_707 and Aq\_880 were

## Significance

**RNase P is a tRNA-processing enzyme of unique architectural diversity: either a catalytic RNA plus one or more (up to 10) proteins, or one (or three) unrelated proteins only. We identified yet another enzyme form in the bacterium *Aquifex aeolicus*, a 23-kDa protein and the smallest known form of RNase P. Apparently, it was acquired by horizontal gene transfer from Archaea. In some other bacteria and many archaea, it is simultaneously present with the presumably more ancient RNA-based enzyme form. Bacteria with both activities may represent the missing link of RNase P evolution, a transition state that had also been once traversed by the *Aquificaceae*, which, however, later lost their RNA-based RNase P.**

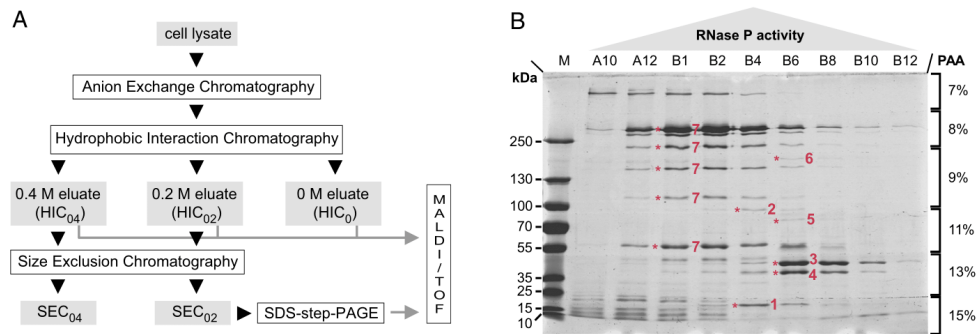
Author contributions: A.I.N., N.B.W., M.G., W.R., and R.K.H. designed research; A.I.N., N.B.W., M.G., and U.T. performed research; U.T. contributed new reagents/analytic tools; A.I.N., N.B.W., M.G., M.L., W.R., and R.K.H. analyzed data; M.G., M.L., W.R., and R.K.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. Email: roland.hartmann@staff.uni-marburg.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1707862114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1707862114/-DCSupplemental).



**Fig. 1.** Partial purification and identification of *A. aeolicus* RNase P. (A) Schematic overview of the purification procedure. The applied chromatography steps and methods are shown as open boxes; the cell lysate and column fractions with RNase P activity are indicated as gray boxes. (B) SDS/PAGE analysis of "SEC<sub>0.2</sub>" fractions with RNase P activity eluting from a SEC column that had been loaded with a "HIC<sub>0.2</sub>" sample [material eluted from the hydrophobic interaction column at 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>]. Fractions with low (A10, B12), increasing (A12, B1, B2), maximum (B4), and decreasing (B6, B8, B10) RNase P activity were loaded in order of their elution from the column (SI Appendix, Fig. S5). Prominent protein bands correlating with RNase P activity were excised and identified by mass spectrometry as: 1, hypothetical protein Aq\_880; 2, polynucleotide phosphorylase (PNPase); 3, uncharacterized protein homologous to RNA pseudouridine synthase from *B. subtilis*; 4, hypothetical protein Aq\_707; 5, Aq\_1754/Aq\_707; 6, Aq\_808/Aq\_707; 7, glutamine synthetase; only two of these proteins (1 and 2) were most abundant in fraction B4, which displayed the highest RNase P activity; M, molecular mass marker.

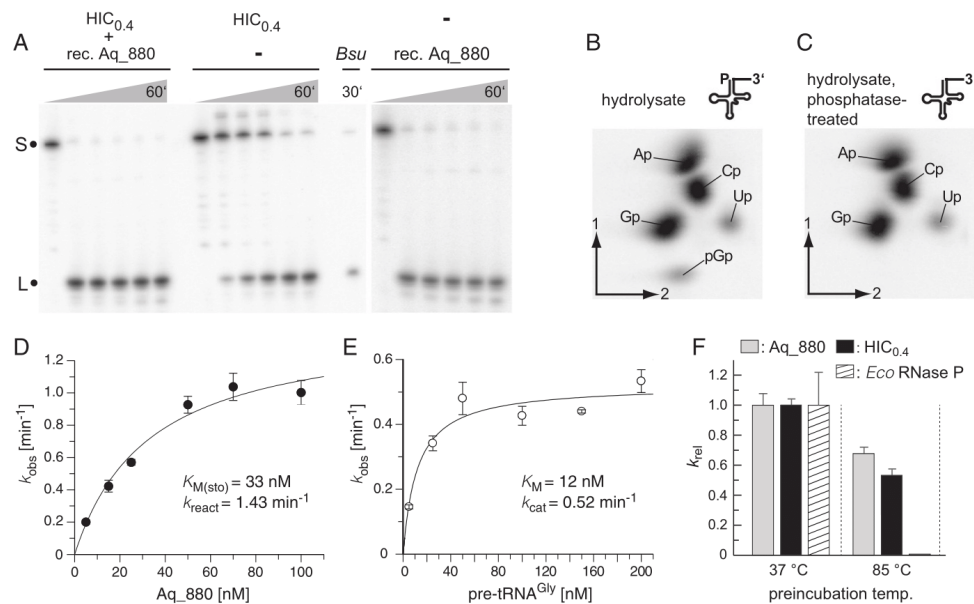
identified in band 6 (Fig. 1B) excised from SDS gels, we tested whether these recombinantly produced *A. aeolicus* proteins support RNase P activity (SI Appendix, Fig. S13). This was done by adding excess amounts of recombinant Aq\_707 or Aq\_221 to a HIC<sub>0.4</sub> fraction and analyzing processing at a pretRNA concentration of 33 nM (SI Appendix, Fig. S13A). The rationale was that Aq\_707 or Aq\_221, if being part of the RNase P enzyme complex and contributing to activity, were likely present in substoichiometric amounts owing to depletion during purification, such that their exogenous addition might boost RNase P activity. However, neither Aq\_707 nor PNPase showed intrinsic endonuclease activity nor did they stimulate the RNase P activity of Aq\_880; PNPase caused some unspecific degradation of the pretRNA and the 5'-leader cleavage product, which we attribute to its 3'-exonuclease activity (SI Appendix, Fig. S13A). To further examine the possibility that Aq\_707, identified together with Aq\_880 in band 6 (Fig. 1B) by MS analysis, might support interaction of Aq\_880 with pretRNA substrates, we analyzed processing under dilute multiple-turnover conditions where substrate binding may be rate-limiting, using 1 nM recombinant Aq\_880, 5 nM pretRNA, and 1 nM recombinant Aq\_707. However, the rate constant of cleavage was essentially identical in the presence and absence of Aq\_707 (SI Appendix, Fig. S13B). These findings argue against a direct interaction of the two proteins. Band 6 in Fig. 1B migrated between the 130- and 250-kDa marker proteins. Protein monomers of Aq\_880 and Aq\_707 have sizes of 23 and 40 kDa, respectively, which left open the possibility that one or two other macromolecules mediate incorporation of the two proteins into a larger complex. However, MS analysis did not detect any other protein in band 6. Finally, addition of total RNA prepared from *A. aeolicus* did not affect the reaction catalyzed by recombinant Aq\_880 (SI Appendix, Fig. S14). We thus concluded that Aq\_880 is the principal constituent of *A. aeolicus* RNase P.

The single-turnover kinetic parameters of recombinant Aq\_880 in a low-salt buffer with 4.5 mM Mg<sup>2+</sup> at 37 °C (Fig. 2D) were determined as  $k_{\text{react}} = 1.43 \text{ min}^{-1}$  and  $K_{\text{M(sto)}} = 33 \text{ nM}$ , thus similar to the values previously obtained for protein-only RNase P (PRORP3) from *Arabidopsis thaliana* [ $k_{\text{react}} = 1.7 \text{ min}^{-1}$ ;  $K_{\text{M(sto)}} \sim 5 \text{ nM}$ ] (15) and similar to activities determined by others (16). The multiple-turnover kinetic parameters determined under the same conditions at 1 nM enzyme and 5–200

nM substrate were  $0.5 \text{ min}^{-1}$  for  $k_{\text{cat}}$  and 12 nM for  $K_{\text{m}}$  (Fig. 2E). The two- to threefold decrease of both parameters relative to those of the single turnover might be explained by product release limiting the rate of the multiple-turnover reaction. SEC indicated that native and recombinant Aq\_880 elute as a ~420-kDa complex (SI Appendix, Fig. S15), suggesting that the protein forms homooligomers (e.g., three hexamers or six trimers). This is in line with a weak band of ~70 kDa observed for recombinant Aq\_880 in SDS gels in addition to the monomer (~23 kDa; Fig. 3B and SI Appendix, Fig. S16). Mass spectrometry identified this additional band at ~70 kDa as Aq\_880 (SI Appendix, Fig. S16), supporting the notion that Aq\_880 has the potential to form stable homotrimers.

We further investigated the thermostability of recombinant Aq\_880 as the hyperthermophile *A. aeolicus* naturally thrives at 85–95 °C. This was done by preincubating recombinant Aq\_880 for 10 min at 85 °C before conducting the cleavage assay at 37 °C. The *E. coli* RNase P holoenzyme served as a control treated in the same way. Preincubation at 85 °C preserved the bulk of Aq\_880 activity, but essentially inactivated *E. coli* RNase P (Fig. 2F, Left). Similar results were obtained when we used a native HIC<sub>0.4</sub> fraction instead of recombinant Aq\_880 (Fig. 2F, Right). These findings confirm the expected thermostability of Aq\_880 and do not provide evidence for substantial differences in thermostability between the native and recombinant Aq\_880.

Bioinformatic analyses predicted limited sequence and structure similarities between Aq\_880 and the PIN (PiT N-terminal) ribonuclease domain. A PIN domain-like fold is also characteristic for the NYN domains of eukaryotic protein-only RNase P (PRORP) enzymes, for which a two metal-ion catalytic mechanism with metal ions coordinated by conserved aspartate residues was proposed (17). Limited similarity of Aq\_880 to the metallo-nuclease domain of *A. thaliana* PRORP isoenzymes was evident from the alignment of the full-length sequences in SI Appendix, Fig. S17A. This alignment also suggested sporadic similarity to the PPR RNA binding and the central domains of PRORPs. Considering that an N-terminal truncation of 8 of the 11 helices in the PPR domain essentially abolished pretRNA-processing activity (17), the N-proximal part of Aq\_880 may exert a corresponding function in pretRNA binding, although this is unclear at present and has to be addressed in future studies. The partial alignment of the major



**Fig. 2.** (A) Processing of *Thermus thermophilus* pre-tRNA<sup>Gly</sup> by the "HIC<sub>0.4</sub>" fraction [material eluted from a HIC column at 0.4 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; Center], by recombinant Aq\_880 (final concentration 0.1 μg/μL) (Right), or by a mixture of both (Left); Bsu, *B. subtilis* RNase P as positive control; S, the 5'-[<sup>32</sup>P]-end-labeled pre-tRNA<sup>Gly</sup> substrate; L, the 5'-leader cleavage product. (B and C) Analysis of the 5'-end of pre-tRNA<sup>Gly</sup> as processed by recombinant Aq\_880 in vitro. Pre-tRNA<sup>Gly</sup> was labeled by in vitro transcription in the presence of [<sup>32</sup>P]GTP and processed with recombinant Aq\_880 under standard conditions to nearly completeness. The 5'-mature tRNA product was gel-purified, and a part of the eluate was treated with alkaline phosphatase. Untreated (B) and phosphatase-treated (C) RNAs were subjected to alkaline hydrolysis, and monophosphate and diphosphate nucleosides were resolved by 2D TLC. The pGp bisphosphate (B) is sensitive to phosphatase pretreatment (C) consistent with its origin from the tRNA's 5'-end. Spot intensities correspond to the number of the respective nucleosides that are followed by a G in the tRNA sequence and, thereby, are radioactively labeled (A, 8; C, 8; G, 9; U, 1). (D) Single-turnover kinetics of recombinant Aq\_880 for the processing of pre-tRNA<sup>Gly</sup> (data points are based on three to six independent determinations for each concentration of Aq\_880); error bars are SDs. Kinetic constants and the standard errors of the curve fit are  $k_{react} = 1.43 \pm 0.14 \text{ min}^{-1}$  and  $K_{M(sto)} = 33 \pm 8 \text{ nM}$ . (E) Multiple-turnover kinetics of pre-tRNA<sup>Gly</sup> processing using 1 nM Aq\_880. Results are based on three independent determinations for each pre-tRNA concentration. Error bars are SDs of the mean. Kinetic constants and the standard errors of the curve fit are  $k_{cat} = 0.52 \pm 0.04 \text{ min}^{-1}$  and  $K_M = 12 \pm 4 \text{ nM}$ . (F) Thermostability of recombinant Aq\_880 (gray bars) or a HIC<sub>0.4</sub> fraction containing native RNase P (black bars) in comparison with the *E. coli* RNase P holoenzyme (white and hatched bars, respectively). RNase P activities were preincubated either at 37 °C or 85 °C before their catalytic activity was determined under single-turnover conditions at 37 °C. Error bars are standard deviations of the mean based on at least three experiments. Note that the activity of *E. coli* RNase P after preincubation at 85 °C was so low that the corresponding bar to the right of the black HIC<sub>0.4</sub> bar does not rise above the baseline.

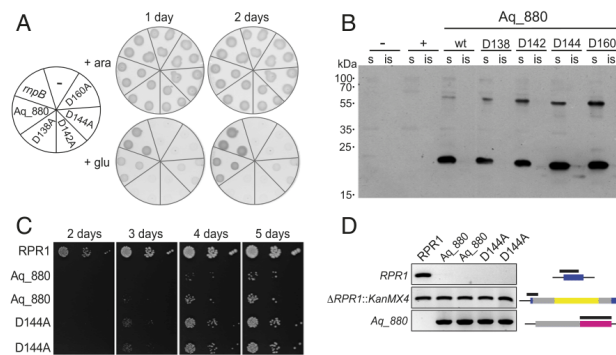
portion of the PRORP metallo-nuclease domain with Aq\_880 (SI Appendix, Fig. S17B) suggested that aspartate residues D138, D142, and D160 of Aq\_880 are functionally equivalent to D399, D475, and D493 in the catalytic center of PRORP1 from *A. thaliana* (SI Appendix, Fig. S18). We also considered a role for D144 of Aq\_880 (corresponding to Y477 in PRORP1) since this residue, assumed to be in close proximity to the putative active site, was found to be strongly conserved (as D or E) in Aq\_880 homologs identified by bioinformatic screens (SI Appendix, Figs. S23 and S24). We mutated the conserved aspartate residues in Aq\_880 individually to alanines and analyzed activity of the resulting recombinant Aq\_880 variants under single-turnover conditions at 37 °C (SI Appendix, Figs. S19 and S20). Mutations D138A, D142A, and D160A essentially abolished activity, while the D144A mutant was fully active.

Aq\_880 (and less efficiently also the D144A variant) was also able to rescue the growth of the conditionally lethal *E. coli* RNase P mutant strain BW (18) (Fig. 3A), demonstrating that the small protein from *A. aeolicus* provides sufficient RNase P activity to

support growth of the *E. coli* host, even if the growth rate was slower than upon complementation with the homologous RNase P RNA gene (*mpB*). Bacteria harboring an empty plasmid or expressing the inactive variants D138A, D142A, and D160A were unable to grow under the same conditions. Western blotting confirmed expression of Aq\_880 mainly in the soluble protein fraction of *E. coli* lysates (Fig. 3B).

Aq\_880 was even able to replace the complex RNase P of *Saccharomyces cerevisiae* (Fig. 3C), a 400-kDa ribonucleoprotein (RNP) consisting of one RNA and nine protein subunits. For this analysis, we employed a previously established plasmid-shuffle protocol (19). Here, the otherwise lethal deletion of the yeast RNase P RNA gene *RPR1* was rescued by the plasmid-based expression of *A. aeolicus* Aq\_880, demonstrating that the highly dissimilar RNase P enzymes are basically exchangeable. The catalytically inactive variant D160A did not give rise to any viable colonies. Yeast cells dependent on Aq\_880 nevertheless grew considerably slower than the parental wild-type strain, with only a slight improvement by the D144A variant (Fig. 3C); the deletion of





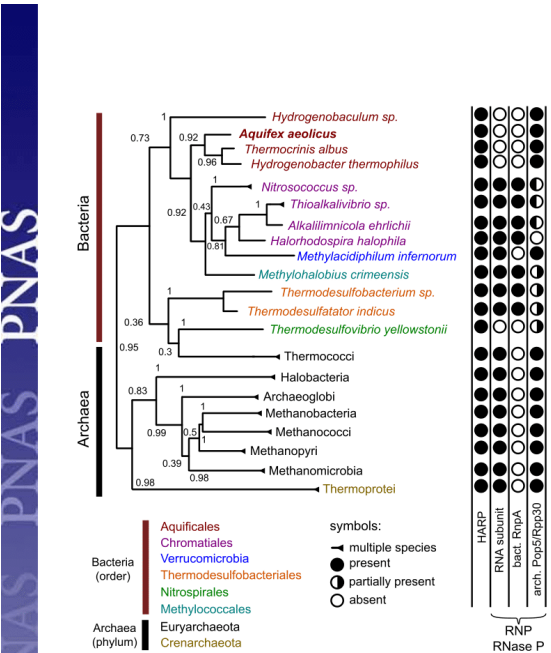
**Fig. 3.** Complementation of *E. coli* and yeast RNase P by Aq\_880. (A) *A. aeolicus* Aq\_880 is able to functionally replace bacterial RNase P in vivo. *E. coli* BW cells transformed with expression plasmids for Aq\_880 or its variants D138A, D142A, D144A, and D160A were grown at 37 °C under permissive (in the presence of arabinose; +ara) or nonpermissive (in the presence of glucose; +glu) conditions. Already after 1 d, cell growth was detectable in *E. coli* BW cells containing the expression vector for Aq\_880; whereas no complementation was observed with the aspartate to alanine mutants D138A, D142A, and D160A, weak colony formation was seen with D144A. *rnpB*: *E. coli* RNase P as positive control. (–), the empty vector as negative control. (B) Expression and solubility of Aq\_880 variants in *E. coli* BW. Soluble (s) and insoluble (is) protein fractions from *E. coli* BW expressing the Aq\_880 variants were analyzed by Western blotting using a polyclonal antiserum to Aq\_880. No signal was obtained when *E. coli* BW was transformed with the empty vector (–) or *E. coli* *rnpB* (+), respectively; molecular mass marker (in kilodaltons) indicated on the left. (C) Aq\_880 and its variant D144A rescue growth upon deletion of the yeast nuclear RNase P RNA gene *RPR1*. Two colonies each obtained through rescue by Aq\_880 or its variant D144A were applied as spots in 10-fold serial dilution to YPD plates, and the growth of the strains was monitored in parallel to a control (rescue by *RPR1*). Note the later appearance and smaller colony size of Aq\_880 and D144A strains indicating slow growth. (D) Genotyping of Aq\_880 and D144A colonies derived from plasmid shuffle. The analysis of a control (*rpr1Δ::kanMX4* [*RPR1*]) and two complementation isolates each (Aq\_880, genotype *rpr1Δ::kanMX4* [Aq\_880]; D144A, genotype *rpr1Δ::kanMX4* [Aq\_880<sup>D144A</sup>]) is shown. The deletion of *RPR1*, the integrity of the chromosomal gene disruption, and the presence of Aq\_880 were verified by PCR. The part of the gene interrogated by the genotyping PCR is indicated by a black bar on top of the gene cartoon (*RPR1*, blue; *kanMX4*, yellow; Aq\_880, magenta; promoter/terminator regions, gray) to the right of each agarose gel panel.

the endogenous *RPR1* gene (*rpr1Δ::kanMX4*) and the presence/absence of plasmid-encoded *RPR1* and Aq\_880 were verified by PCR (Fig. 3D). The poor growth contrasts with similar complementation experiments using eukaryal PRORPs, of which some, but not all, resulted in strains with wild-type-like growth properties (19). However, even strains harboring the less effectively complementing *A. thaliana* PRORP1 grew better than any of the Aq\_880 strains. The slow growth could be due to problems in expressing sufficient bacterial protein in the eukaryal host, unfavorable subcellular localization, differences in substrate recognition between Aq\_880 and yeast nuclear RNase P, or a combination of them leading to retarded biogenesis of mature tRNAs in general, or of a subset of tRNAs as observed for *A. thaliana* PRORP1 complementation in *E. coli* (20).

Bioinformatic screens identified numerous Aq\_880 homologs in Archaea, few in Bacteria (Fig. 4 and SI Appendix, Figs. S22–S24 and Table S1), but none in Eukarya. We therefore propose to name this type of protein-only RNase P “HARP” for Homolog of *Aquifex* RNase P. The HARP proteins fall within the PIN\_5 group of a recently published classification of the PIN domain-like superfamily (21). The relatively large group of archaeal organisms that encode HARP (SI Appendix, Table S1) have in common that they additionally encode an RNP RNase P. This suggests at least partially divergent functions of HARP and RNP enzymes in these Archaea, rather than reflecting evolutionary transition states where RNP RNase P is about to be replaced by HARP. In the latter scenario, one would have expected an irregular distribution of the two RNase P types among these archaeal species. The group of Bacteria-encoding HARPs was considerably smaller than the archaeal group (SI Appendix, Table S1). In the *Aquificaceae*, which all lack genes for a canonical bacterial RNase P, we identified homologs of the Archaea-specific RNase P protein subunits Pop5 and Rpp30, which bind to the catalytic domain of archaeal

RNase P RNA (22). This suggests that the genes for HARP, Pop5, and Rpp30 were acquired by horizontal gene transfer from Archaea, in line with bioinformatic evidence that *A. aeolicus* has acquired at least 10% of its protein-coding genes by horizontal gene transfer from archaeal organisms (23). In the genomes of the *Aquificaceae* *A. aeolicus* and *Hydrogenobacter thermophilus* TK6, HARP and Pop5 genes are separated by only 12.5 and 3.1 kb, respectively, which may be evidence for their horizontal co-transfer. The findings may also suggest a functional link between HARP and Pop5/Rpp30. *Thermodesulfobacterium yellowstonii* seems to represent a situation similar to the *Aquificaceae* (no RNP RNase P), although a Pop5 homolog could not be identified (SI Appendix, Table S1). Remarkably, we identified a few bacteria that encode both an RNP RNase P and HARP. In these cases, either Rpp30 or Pop5 is not detectable (SI Appendix, Table S1), suggesting relaxed constraints to keep Rpp30 and Pop5 when RNA-based RNase P is expressed. All *mpA* genes of this bacterial group with assumed dual RNase P activities encode regular RnpA proteins. Likewise, their *mpB* genes encode RNase P RNAs that largely conform to canonical bacterial type A structures. This raises the question whether HARP and RNP RNase P both exert RNase P activity when coexpressed in the same organism. So far, evidence suggesting that both enzyme types might have RNase P activity within the same organism was obtained for one bacterium, *Thermodesulfatator indicus*, and one archaeon *Methanothermobacter thermautotrophicus*. *T. indicus* RNase P RNA has RNase P activity in vitro (Table 1), and the RNase P activity of the *M. thermautotrophicus* RNP enzyme has been thoroughly characterized (24–26). The HARP enzymes of both organisms were active in vitro and able to complement *E. coli* BW bacteria in vivo (Table 1 and SI Appendix, Fig. S21).

We finally directed our attention to Archaea associated with deviations from canonical RNase P processing. The archaeal symbiont



**Fig. 4.** Phylogenetic distribution of HARP. Condensed phylogenetic tree of HARPs. Bacteria were condensed at order level, Archaea at phylum level. Bootstrap values are indicated at the branches (inferred from 1,000 replicates). Filled and empty circles indicate the presence or absence, respectively, of RNase P RNA (RNA subunit), the bacterial RNase P protein (RnpA), or the archaeal RNase P protein subunits Pop5/Rpp30; half-filled circles indicate that we identified for (at least some) bacterial species in this order only an Rpp30 (white/black circles) or Pop5 homolog (black/white circle), but not both.

*Nanoarchaeum equitans* lacks RNase P and the need for tRNA 5'-end maturation is obviated by transcription of leaderless tRNAs (27). In the archaeon *Ignicoccus hospitalis* KIN4 I, which is the host for *N. equitans* (28), both HARP and RNA-based RNase P components were identified (SI Appendix, Table S1), although Pop5 and Rpp30 homologs were only found upon increasing the *e* value from  $10^{-10}$  to  $10^{-6}$  in genome searches. *Pyrobaculum*, *Caldivirga*, and *Vulcanisaeta* species encode shortened RNase P RNAs lacking the specificity domain (22). This suggested the possibility that these minimized enzymes have restricted functionality, which might be compensated by the presence of a HARP enzyme. However, even when increasing the *e* value from  $10^{-10}$  to  $10^{-4}$  in genome searches, no HARP candidates could be identified in any of these Archaea.

Conclusions

We discovered a minimal and RNA-free RNase P in the hyperthermophilic bacterium *A. aeolicus*, which was apparently acquired by horizontal gene transfer from an archaeon. Thus, we not only solved the last cold case in the RNase P field, but identified a previously unknown class of RNase P. Homologs were found in only a few Bacteria but many Archaea. At least for one archaeon and one bacterium analyzed so far we could provide evidence that both forms, RNA-based and protein-only RNase P, catalyze tRNA 5'-end maturation. As these findings are based on in vitro data and some complementation results in *E. coli* (Table 1), future studies will have to clarify if these homologs indeed function as genuine RNase P enzymes in vivo. Our findings have revealed yet another face of the unique architectural diversity

of the RNase P enzyme family, comprised of diverse members of RNA and protein catalysts. RNA-based and protein-only RNase P enzymes have also been found in Eukarya (6), but appear to be mutually exclusive there, i.e., in no case they coexist within the same genetic compartment. With the identification of Bacteria that have protein- and RNA-based RNase P, we seem to have identified the kind of evolutionary transition state that was traversed by the *Aquificaceae*, which then, at some point in evolution, lost their "ancient" RNA-based enzyme. The situation in Archaea is evolutionarily and biologically intriguing as well, as the presence of this protein-based enzyme form seems to be linked to the presence of an RNA-based RNase P, suggesting that the two enzyme types functionally complement each other in tRNA processing, or one may exert another function, unrelated to tRNA processing, in Archaea. Finally, it will be challenging to understand why the *Aquificaceae* acquired and have retained the two archaeal RNase P RNA binding proteins Pop5 and Rpp30.

Materials and Methods

**Growth of Bacterial Strains.** *A. aeolicus* VF5 cell pellets were purchased from M. Thomm and R. Huber (University of Regensburg, Germany). After growing for 1 d, cells were harvested as described (29) in the late exponential phase. *E. coli* strains were generally grown in liquid or on solid LB medium (10 g of tryptone, 5 g of yeast extract, 10 g of sodium chloride and, if required, 15 g of agar-agar per liter) at 37 °C under shaking (220 rpm; warm air incubation shaker, GFL 3033) in the case of liquid cultures. For the selection of strains expressing antibiotic resistance genes, growth media were supplemented with ampicillin (100 µg/mL), kanamycin (50 µg/mL), and/or chloramphenicol (34 µg/mL). *E. coli* DH5α was used as the host strain for cloning, *E. coli* Rosetta (DE3) for protein expression, and *E. coli* BW (18) for complementation studies. Liquid overnight cultures were usually incubated in 2–8 mL of LB medium containing the appropriate antibiotics and, if required, specific carbon sources.

**Purification and Enrichment of RNase P Activity.** Purification of *A. aeolicus* RNase P by AEC, HIC, and SEC, as well as size exclusion chromatography to study the oligomeric state of Aq\_880, was performed with an ÄKTA basic 10 (GE Healthcare) FPLC system at room temperature. Details are described in SI Appendix.

**Step Gradient SDS/PAGE and Qualitative and Quantitative Mass Spectrometry Analysis.** The procedures are detailed in SI Appendix.

**In Vitro Transcription of RNAs, 5'-[<sup>32</sup>P]-Endlabeling, RNase P Processing Assays.** This was done according to standard protocols (30) detailed in SI Appendix.

**Pretreatment with Micrococcal Nuclease.** RNase P enzyme solutions were preincubated with micrococcal nuclease (Thermo Scientific) before addition

**Table 1.** Activity of RNP RNase P and HARP in selected bacteria and archaea

Organism	RNP RNase P in vitro	HARP	
		in vitro	in vivo
Bacteria			
<i>E. coli</i>	+++	—	—
<i>A. aeolicus</i>	—	+++	+++
<i>T. indicus</i>	+	++	+++
Archaeon			
<i>M. thermotrophicus</i>	Active*	+	++

RNP RNase P (RNA alone and holoenzyme) and HARP in vitro activities were tested as described in SI Appendix. —, enzyme not present; +(+), relative enzyme activities in processing assays or relative efficiencies in genetic complementation of strain BW; for example, in RNA-alone reactions, the  $k_{\text{obs}}$  of pre-tRNA<sup>Gly</sup> cleavage was 0.54 min<sup>−1</sup> (++) for *E. coli* and 0.11 min<sup>−1</sup> (+) for *T. indicus* RNase P RNA; in complementation experiments (HARP in vivo), +++ was assigned to colony densities roughly corresponding to those obtained for BW bacteria complemented with aq\_880 after overnight incubation at 37 °C.

\*Shown to be active in previous studies (24, 25).

of substrate and excess amounts of carrier RNA (6S-1 RNA from *B. subtilis*). The experimental details can be found in [SI Appendix](#).

**Cloning Procedures, Recombinant Protein Expression, and Complementation Studies in *E. coli* and Yeast.** This was performed by standard procedures; complementation analyses were carried out as described (18, 19). Experimental details are provided in [SI Appendix](#).

**Determination of Single and Multiple Turnover Kinetic Parameters for pre-tRNA<sup>Gly</sup> Processing by Aq\_880.** The experimental details are described in [SI Appendix](#).

**5'-End Group Analysis.** 5'-end group analysis of mature tRNA<sup>Gly</sup> derived from processing of pre-tRNA<sup>Gly</sup> by Aq\_880 was performed as described (31).

**Bioinformatic Prediction of RNase P Components.** Based on the Aq\_880 reference sequence, Aq\_880 homologs were searched in two iterations in all bacterial and archaeal genomes provided at the National Center for Biotechnology Information (NCBI) (download 2015-10-06) using blastp+ at the proteome level and tblastn+ at the genome level with an *e* value of 10<sup>-10</sup> (32, 33). Similarly, we used the reference sets provided in ref. 34 for Rpp29 (Pop4) and a self-assembled set of well-annotated reference sequences for RnpA homolog searches (the bacterial P protein, also termed C5). The latter comprised sequences from *Avibacterium*, *Bibersteinia*, *Enterobacter*, *Escherichia*, *Herbaspirillum*, *Mycobacterium*, and *Serratia*. The RNA subunit was predicted using Infernal (35) with the Rfam v12 models RF00010, RF00011, and RF00373 (bacterial RNase P class A, B, and archaeal RNase P, respectively) (36). In the case of multiple hits, only the best one was chosen. All results were aligned using Clustal Omega (37) and manually inspected for likely false positives (none identified). The alignment of Aq\_880 homologs ([SI Appendix, Fig. S23](#))

and the reference sequences/IDs used for the RnpA search can be downloaded at [bioinf.pharmazie.uni-marburg.de/supplements/harp\\_2016/](http://bioinf.pharmazie.uni-marburg.de/supplements/harp_2016/). Based on the alignment region 19–253, a WebLogo ([SI Appendix, Fig. S24](#)) was generated (38). An overview of the results is given in [SI Appendix, Table S1](#).

**Phylogenetic Tree.** The phylogenetic tree of Aq\_880 homologs was generated using all identified sequences (see above) with a RAxML (39) rapid bootstrap analysis with 1,000 bootstrap runs to search for the best-scoring maximum likelihood tree with respect to the GTR substitution model and the Gamma model of rate heterogeneity (40). For the sake of clarity, bacteria were condensed at order level, archaea at phylum level for Fig. 4. [SI Appendix, Table S1](#) shows the results in more detail.

**Alignment of Aq\_880 and *A. thaliana* PRORP1-3.** The alignment of Aq\_880 to the metallo-nuclease domains of *A. thaliana* PRORP1-3 was performed in a semi-automatic manner. The NYN domains were manually aligned based on the NYN WebLogo provided in ref. 4. The regions upstream and downstream of the NYN domain were aligned using Clustal Omega (37). NCBI accession codes were NP\_850186 for *A. thaliana* PRORP1, NP\_179256 for PRORP2, and NP\_193921 for PRORP3.

**ACKNOWLEDGMENTS.** We thank Diogo Monteiro for his help with the yeast complementation experiments and Dominik Helmecke for support with respect to processing assays, Johann Heider and Wolfgang Buckel for giving us access to the French press and ultracentrifuge, and Tina Krieg for technical assistance in sample preparation for mass spectrometry. The Superose 6 column was a loan from Peter Friedhoff, Justus-Liebig-Universität Gießen. *A. aeolicus* ribosomal subunits were kindly prepared in the laboratory of Ciaran Condon (Paris).

- Hartmann E, Hartmann RK (2003) The enigma of ribonuclease P evolution. *Trends Genet* 19:561–569.
- Klemm BP, et al. (2016) The diversity of ribonuclease P: Protein and RNA catalysts with analogous biological functions. *Biomolecules* 6:E27.
- Holzmann J, et al. (2008) RNase P without RNA: Identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* 135:462–474.
- Gober A, et al. (2010) A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol* 17:740–744.
- Taschner A, et al. (2012) Nuclear RNase P of *Trypanosoma brucei*: A single protein in place of the multicomponent RNA-protein complex. *Cell Rep* 2:19–25.
- Lechner M, et al. (2015) Distribution of ribonucleoprotein and protein-only RNase P in Eukarya. *Mol Biol Evol* 32:3186–3193.
- Hartmann RK, Gössinger M, Späth B, Fischer S, Marchfelder A (2009) The making of tRNAs and more – RNase P and tRNase Z. *Prog Mol Biol Transl Sci* 85:319–368.
- Swanson RV (2001) Genome of *Aquifex aeolicus*. *Methods Enzymol* 330:158–169.
- Willkomm DK, Feltens R, Hartmann RK (2002) tRNA maturation in *Aquifex aeolicus*. *Biochimie* 84:713–722.
- Li Y, Altman S (2004) In search of RNase P RNA from microbial genomes. *RNA* 10: 1533–1540.
- Marszałkowski M, Willkomm DK, Hartmann RK (2008) 5'-end maturation of tRNA in *aquifex aeolicus*. *Biol Chem* 389:395–403.
- Lechner M, et al. (2014) Genomewide comparison and novel ncRNAs of *Aquificales*. *BMC Genomics* 15:522.
- Lombo TB, Kaberdin VR (2008) RNA processing in *Aquifex aeolicus* involves RNase E/G and an RNase P-like activity. *Biochem Biophys Res Commun* 366:457–463.
- Barrera A, Pan T (2004) Interaction of the *Bacillus subtilis* RNase P with the 30S ribosomal subunit. *RNA* 10:482–492.
- Brillante N, et al. (2016) Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P. *Nucleic Acids Res* 44:2323–2336.
- Howard MJ, et al. (2016) Differential substrate recognition by isozymes of plant protein-only Ribonuclease P. *RNA* 22:782–792.
- Howard MJ, Lim WH, Fierke CA, Koutmos M (2012) Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proc Natl Acad Sci USA* 109:16149–16154.
- Wegscheid B, Hartmann RK (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P in vivo. *RNA* 12: 2135–2148.
- Weber C, Hartig A, Hartmann RK, Rossmanith W (2014) Playing RNase P evolution: Swapping the RNA catalyst for a protein reveals functional uniformity of highly divergent enzyme forms. *PLoS Genet* 10:e1004506.
- Göbinger M, et al. (2017) Protein-only RNase P function in *Escherichia coli*: Viability, processing defects and differences between PRORP isoenzymes. *Nucleic Acids Res* 45: 7441–7454.
- Matelska D, Steczkiewicz K, Ginalski K (2017) Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Res* 45:6995–7020.
- Lai LB, et al. (2010) Discovery of a minimal form of RNase P in *Pyrobaculum*. *Proc Natl Acad Sci USA* 107:22493–22498.
- Aravind L, Tatusov RL, Wolf YI, Walker DR, Koonin EV (1998) Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. *Trends Genet* 14: 442–444.
- Andrews AJ, Hall TA, Brown JW (2001) Characterization of RNase P holoenzymes from *Methanococcus jannaschii* and *Methanothermobacter thermoautotrophicus*. *Biol Chem* 382:1171–1177.
- Hall TA, Brown JW (2002) Archaeal RNase P has multiple protein subunits homologous to eukaryotic nuclear RNase P proteins. *RNA* 8:296–306.
- Hall TA, Brown JW (2004) Interactions between RNase P protein subunits in archaea. *Archaea* 1:247–254.
- Randau L, Schröder I, Söll D (2008) Life without RNase P. *Nature* 453:120–123.
- Paper W, et al. (2007) *Ignicoccus hospitalis* sp. nov., the host of 'Nanoarchaeum equitans'. *Int J Syst Evol Microbiol* 57:803–808.
- Huber R, et al. (1992) *Aquifex pyrophilus* gen. nov. sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. *Syst Appl Microbiol* 15: 340–351.
- Busch S, Kirsebom LA, Notbohm H, Hartmann RK (2000) Differential role of the intermolecular base-pairs G292-C(75) and G293-C(74) in the reaction catalyzed by *Escherichia coli* RNase P RNA. *J Mol Biol* 299:941–951.
- Rossmanith W, Tullo A, Potuschak T, Karwan R, Sbisà E (1995) Human mitochondrial tRNA processing. *J Biol Chem* 270:12885–12891.
- Tatusova T, et al. (2015) Update on RefSeq microbial genomes resources. *Nucleic Acids Res* 43:D599–D605.
- Camacho C, et al. (2009) BLAST+: Architecture and applications. *BMC Bioinformatics* 10:421.
- Rosenblad MA, López MD, Piccinelli P, Samuelsson T (2006) Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res* 34:5145–5156.
- Nawrocki EP, Eddy SR (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29:2933–2935.
- Nawrocki EP, et al. (2015) Rfam 12.0: Updates to the RNA families database. *Nucleic Acids Res* 43:D130–D137.
- Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188–1190.
- Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
- Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* 25:1307–1320.



## 4.1.2 Supplementary Informationen

### Supplementary Information (SI) Appendix

#### Minimal and RNA-free RNase P in *Aquifex aeolicus*

Astrid I. Nickel<sup>1</sup>, Nadine B. Wäber<sup>1</sup>, Markus Gößringer<sup>1</sup>, Marcus Lechner<sup>1</sup>, Uwe Linne<sup>2</sup>,  
 Ursula Toth<sup>3</sup>, Walter Rossmanith<sup>3</sup>, Roland K. Hartmann<sup>1†</sup>

### Table of Contents

<b>Three step purification procedure</b>	p.3
Anion exchange chromatography (AEC)	p.3
Hydrophobic interaction chromatography (HIC)	p.4
Size exclusion chromatography (SEC)	p.6
<b>Analysis of partially purified SEC fractions</b>	p.7
<b>Quantitative mass spectrometry (liquid samples with RNase P activity)</b>	p.9
<b>Excluding an <i>E. coli</i> RNase P RNA contamination in recombinant</b>	
<b>Aq_880 preparations</b>	p.11
RT-PCR	p.11
Treatment with micrococcal nuclease (MN)	p.11
<b>Addition of recombinant Aq_707 / Aq_221 (PNPase) to <i>A. aeolicus</i></b>	
<b>fractions with enriched RNase P activity</b>	p.12
<b>Activity of recombinant Aq_880 in the presence of <i>A. aeolicus</i> total RNA</b>	p.14
<b>Thermostability of recombinant Aq_880</b>	p.14
<b>Oligomerization of recombinant Aq_880</b>	p.15
<b>Identification of active site residues in Aq_880</b>	p.17
<b>Bioinformatic analysis</b>	p.20
Evaluation of proteins identified in quantitative and qualitative mass spectrometry analyses	p.20
Analysis of structural domains	p.21
Occurrence of protein candidates within the <i>Aquificales</i> and alignment of bacterial and archaeal Aq_880 homologs	p.22
Table S1	p.25
<b>Supplementary Materials and Methods</b>	p.28
Purification and enrichment of RNase P activity	p.28
Cell lysis using a French press	p.28
Anion exchange chromatography	p.29
Hydrophobic interaction chromatography	p.29
Size exclusion chromatography	p.30
Step gradient SDS-PAGE	p.30
Qualitative and quantitative mass spectrometry analysis	p.31
<i>In vitro</i> transcription of <i>T. thermophilus</i> ptRNA <sup>Gly</sup> , <i>E. coli</i> RNase P RNA and <i>B. subtilis</i> 6S-1 RNA	p.32
Table S2	p.32
Dephosphorylation, 5'-phosphorylation and 5'-[ <sup>32</sup> P]-endlabeling of RNA	p.33
Processing assays	p.33
Partially purified <i>A. aeolicus</i> fractions	p.33

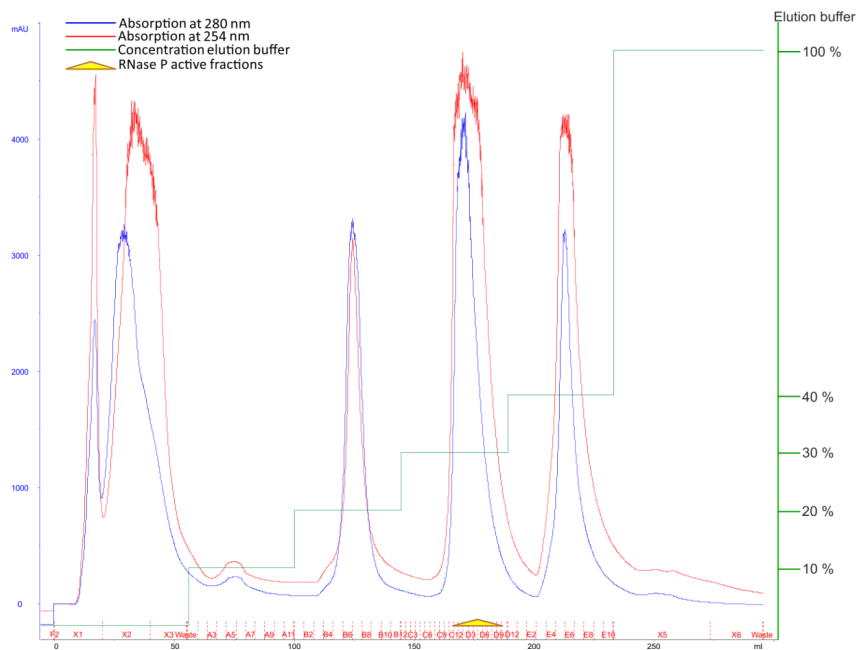


Activity assay with recombinant Aq_880 variants .....	p.34
Determination of single-turnover kinetic parameters for pre-tRNA <sup>Gly</sup> processing by Aq_880.....	p.35
Determination of multiple-turnover kinetic parameters for pre-tRNA <sup>Gly</sup> processing by Aq_880.....	p.35
<i>E. coli</i> and <i>B. subtilis</i> RNase P holoenzyme and <i>E. coli</i> P RNA-alone assays .....	p.35
Thermostability test for Aq_880 and <i>E. coli</i> RNase P .....	p.35
Pretreatment with micrococcal nuclease (MN) .....	p.36
Cell lysate supplemented with recombinant proteins .....	p.36
Aq_880 supplemented with total RNA .....	p.36
Activity assay with other bacterial RNase P RNAs .....	p.37
Activity assays using other HARP proteins .....	p.37
Construction of bacterial expression vectors for Aq_880 .....	p.37
Construction of expression vectors for Aq_880 variants D160A, D138A, D144A and D142A .....	p.38
Table S3 .....	p.38
Expression and purification of recombinant proteins .....	p.39
RT-PCR analysis .....	p.40
<i>A. aeolicus</i> total RNA preparation .....	p.41
Phenol-chloroform extraction .....	p.41
Acetic acid extraction .....	p.41
Size exclusion chromatography of recombinant Aq_880 .....	p.41
Complementation studies .....	p.42
<i>A. aeolicus</i> proteins identified by mass spectrometry - occurrence of homologs in the <i>Aquificales</i> .....	p.42
References .....	p.44

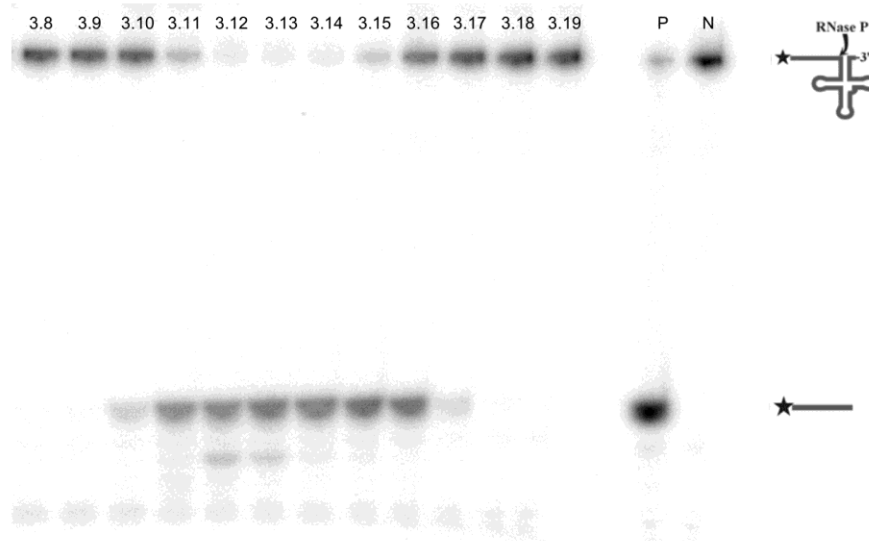
### Three step purification procedure

#### Anion exchange chromatography (AEC)

After cell lysis using a French press and ultracentrifugation, the supernatant was purified by anion exchange chromatography (AEC) using a Diethylaminoethyl (DEAE) Sepharose Fast Flow column (GE Healthcare). RNase P activity eluted in step three at approximately 340 mM ammonium chloride.



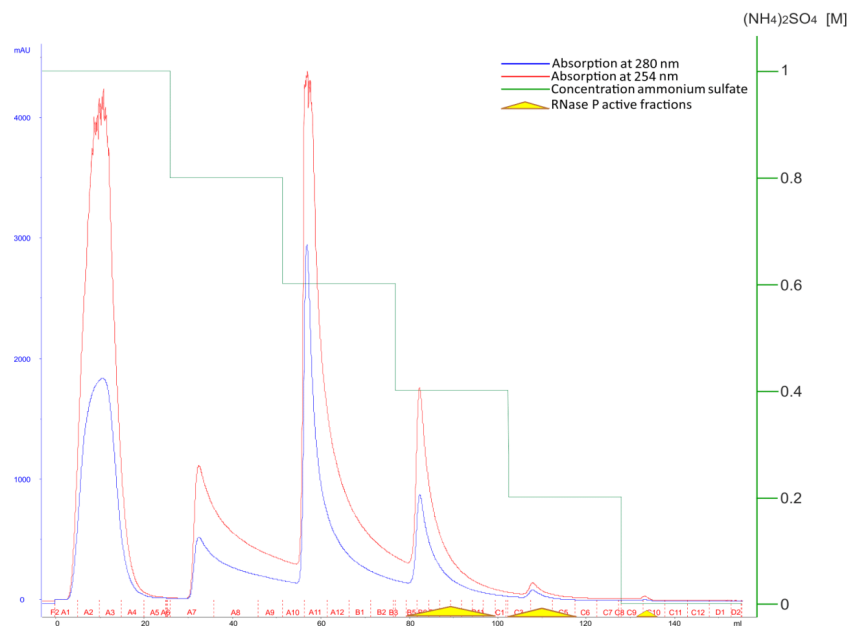
**Figure S1:** AEC (DEAE) chromatography of *A. aeolicus* cell lysates. Purification was performed by applying a step gradient, starting with buffer A60 containing 60 mM ammonium chloride. In each elution step the concentration of elution buffer A1000 containing 1 M  $\text{NH}_4\text{Cl}$  was increased by 10%. Eluted fractions were tested for RNase P activity. Active fractions eluted in step 3 (marked with a yellow triangle) at 30% elution buffer.



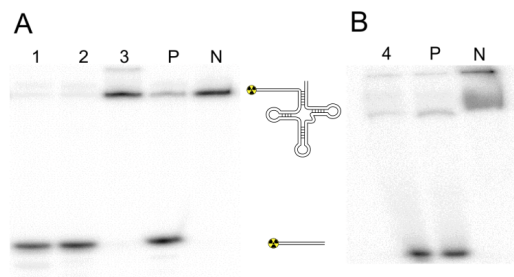
**Figure S2:** DEAE fractions were tested for RNase P activity and analyzed on a 20% denaturing PAA gel containing 8 M urea. 5'-[<sup>32</sup>P]-labeled pre-tRNA<sup>Gly</sup> was incubated with 5 µl of each fraction eluted in step 3 of the AEC. 3.8 – 3.19: fractions collected in step 3. P: positive control, cleavage of pre-tRNA<sup>Gly</sup> by *E. coli* P RNA in the presence of 10 mM Mg(OAc)<sub>2</sub>; N: negative control (ddH<sub>2</sub>O instead of enzyme). Highest RNase P activity was present in fractions 3.12 – 3.14. For assay details, see Suppl. Materials and Methods below, paragraph "Processing assays".

#### *Hydrophobic interaction chromatography (HIC)*

After anion exchange chromatography, DEAE fractions with substantial RNase P activity were pooled and concentrated using Amicon Ultra-15 centrifugal filters. The concentrated fractions were adjusted to buffer A2, which corresponds to buffer A60 but additionally containing 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and loaded onto a Phenyl Sepharose column. Elution was again performed stepwise, starting with buffer A2 and replacing it to 20, 40, 60, 80 and 100% with buffer A60 [no (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>] in elution steps 1 to 5. Fractions with RNase P activity eluted mainly in steps 3 and 4 at 0.4 and 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, respectively. In the following, RNase P fractions eluting from the HIC column at 0.4 M and 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> were denoted as "HIC<sub>0.4</sub>" and "HIC<sub>0.2</sub>", respectively. Only minor RNase P activity was detectable when the HIC<sub>0.4</sub> and HIC<sub>0.2</sub> fractions were tested directly after elution. However, activity could be rescued by subsequent dialysis against buffer A60. For fractions that eluted during HIC in step 5 with no (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> in the elution buffer (fractions denoted as "HIC<sub>0</sub>"), dialysis could be omitted and RNase P activity was detected after pooling and concentration of several HIC<sub>0</sub> fractions, likely due to low levels of RNase P activity in individual fractions from single column runs.



**Figure S3:** Chromatogram of the Phenyl Sepharose (HIC) column. After sample loading, the ammonium sulfate concentration was decreased from 1 to 0 M in 0.2 M steps. Yellow triangles indicate fractions with RNase P activity that eluted at 0.4 M  $(\text{NH}_4)_2\text{SO}_4$  in step 3 ( $\text{HIC}_{0.4}$ ), at 0.2 M  $(\text{NH}_4)_2\text{SO}_4$  in step 4 ( $\text{HIC}_{0.2}$ ) and in the last step without  $(\text{NH}_4)_2\text{SO}_4$  ( $\text{HIC}_0$ ).

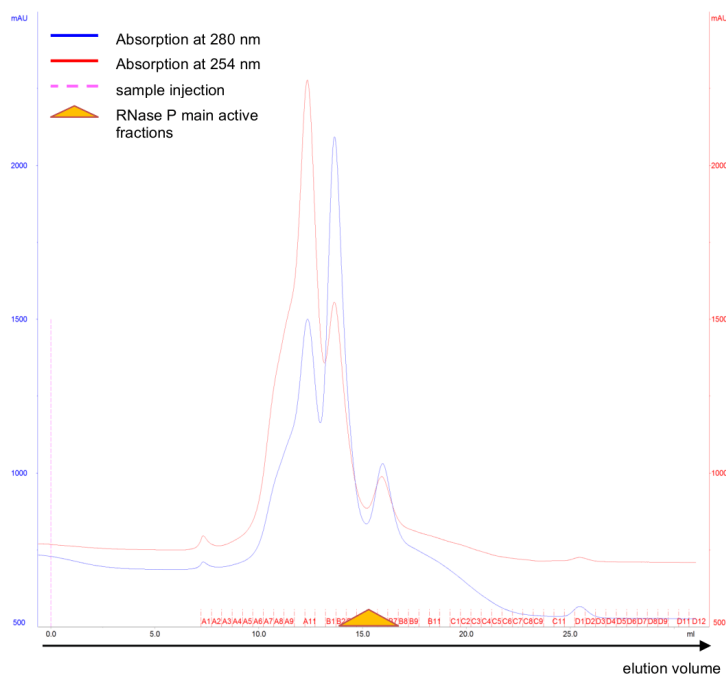


**Figure S4:** (A) RNase P activity test of HIC fractions that eluted at 0.2 M ( $\text{HIC}_{0.2}$ ; lane 1), 0.4 M ( $\text{HIC}_{0.4}$ ; lane 2) and 0.6 M ( $\text{HIC}_{0.6}$ ; lane 3)  $(\text{NH}_4)_2\text{SO}_4$ . For the RNase P activity test, respective fractions of two independent HIC purifications were pooled, dialyzed and concentrated. (B) RNase P activity in the HIC fraction that eluted in step 5 at 0 M  $(\text{NH}_4)_2\text{SO}_4$  ( $\text{HIC}_0$ ; lane 4). For the RNase P activity test, corresponding fractions of 4 independent HIC purifications were pooled and concentrated. P: positive control, cleavage of pre-tRNA<sup>Gly</sup> by *E. coli* P RNA as in Fig. S2; N: negative control (ddH<sub>2</sub>O instead of enzyme).

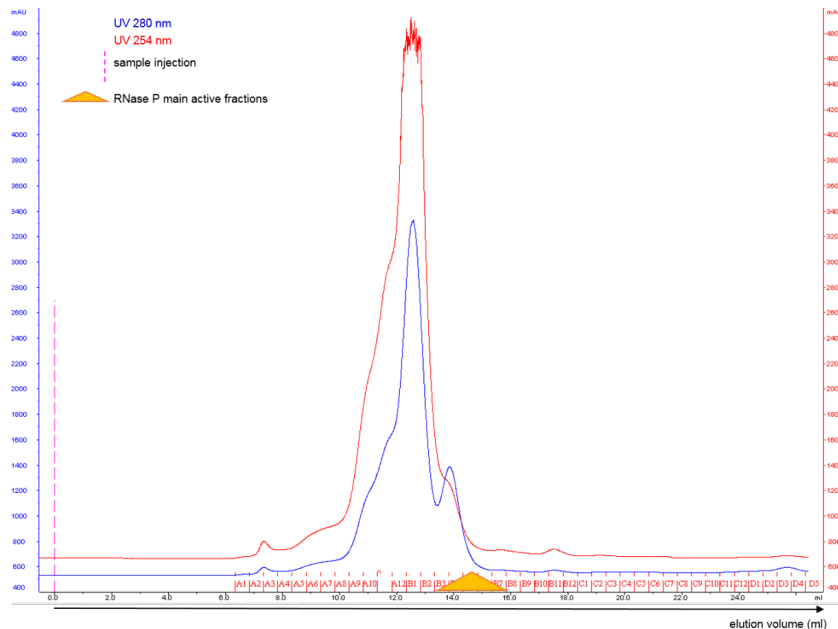
### Size exclusion chromatography (SEC)

In the third purification step, HIC fractions with RNase P activity were purified by size exclusion chromatography (SEC). Since the elution profiles of the fractions HIC<sub>0.4</sub> and HIC<sub>0.2</sub> were different, both fractions were (after dialysis and concentrating) purified separately by size exclusion chromatography using a Superose 6 column. The eluted fractions containing RNase P activity were termed "SEC<sub>0.4</sub>" and "SEC<sub>0.2</sub>", respectively. The separation of fractions with RNase P activity from the bulk of proteins was more efficient when applying the concentrated HIC<sub>0.2</sub> relative to the HIC<sub>0.4</sub> fractions.

### SEC<sub>0.2</sub> chromatogram



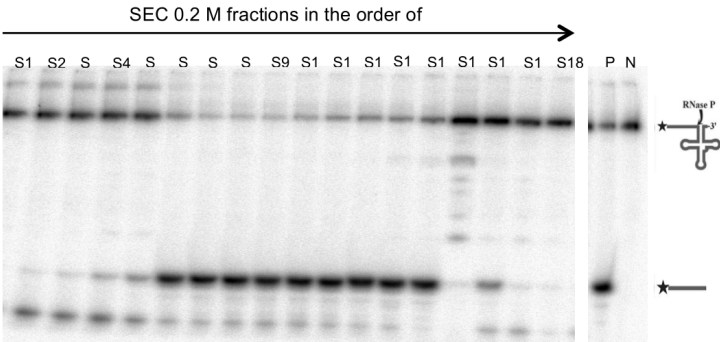
**Figure S5:** Chromatogram of the Superose 6 (SEC) column upon loading of a HIC<sub>0.2</sub> fraction. Fractions containing the bulk of RNase P activity reproducibly eluted after the major absorption peaks (elution buffer: A60), coinciding with a minimum in the UV absorption profile. The elution volume suggested a mass of approximately 350 kDa.

SEC<sub>0.4</sub> chromatogram

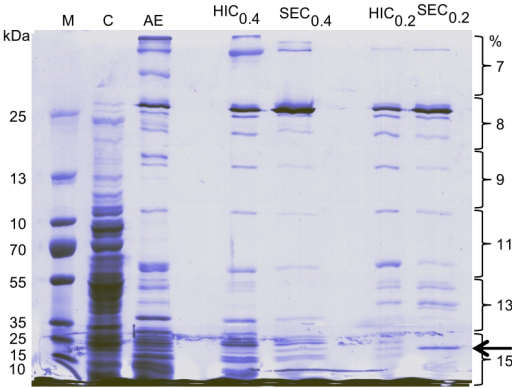
**Figure S6:** Chromatogram of the Superose 6 (SEC) column upon loading of a HIC<sub>0.4</sub> fraction. Fractions containing the bulk of RNase P activity coeluted with the descending shoulder of the second smaller peak (elution buffer: A260). The elution volume of fractions containing the bulk of RNase P activity suggested a mass of approximately 350 kDa.

### Analysis of partially purified SEC fractions

After SEC chromatography of HIC<sub>0.2</sub> samples, the eluted "SEC<sub>0.2</sub>" fractions were tested for RNase P activity and analyzed by 20% denaturing PAGE (Fig. S7). Several elution fractions were selected for further analysis by step gradient SDS-PAGE (Fig. 1B) of the main manuscript), representing the different regions of the RNase P activity peak: (I) fractions with no or weak RNase P activity [A10-A11], (II) fractions with increasing activity [A12, B1, B2], (III) one of the fractions with maximum activity [B4], (IV) fractions with gradually decreasing RNase P activity [B6, B8, B10] and (V) a fraction with low activity [B12].



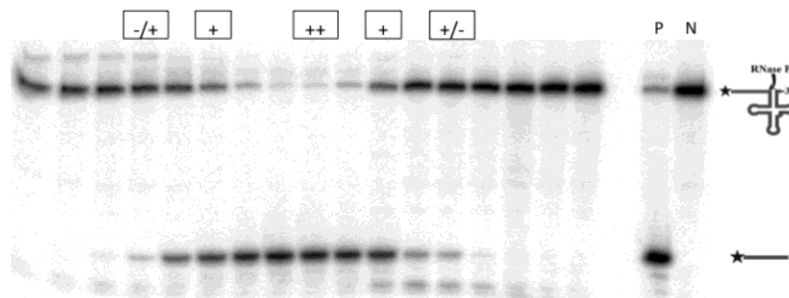
**Figure S7.** SEC<sub>0.2</sub> fractions covering the RNase P activity peak were tested for pre-tRNA processing; samples were analyzed by 20% denaturing PAGE. At the top of the gel, fractions are indicated in the order of elution from the Superose 6 column (S1 – S18); P: positive control, cleavage of pre-tRNA<sup>Gly</sup> by *E. coli* P RNA as in Fig. S2; N: negative control (ddH<sub>2</sub>O instead of enzyme).



**Figure S8.** Protein pattern of RNase P active samples from each purification step. Samples of each purification step were analyzed by step gradient SDS-PAGE (SDS gel with several layers of different PAA concentrations). 2.5 µl of cell lysate (CL) were loaded in the second lane next to the protein marker (M; PageRuler Plus Prestained Protein Ladder, Thermo Scientific). "AEC" depicts 2.5 µl of a concentrated pool of RNase P activity fractions that eluted from the anion exchange column. The samples HIC<sub>0.4</sub> and HIC<sub>0.2</sub> (5 µl each) are the pooled, dialyzed and concentrated RNase P activity fractions that eluted in different HIC steps with elution buffer containing 0.4 M or 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. SEC<sub>0.4</sub> and SEC<sub>0.2</sub> are fractions (15 µl each) showing the highest RNase P activity after purification of the HIC<sub>0.4</sub> and HIC<sub>0.2</sub> samples by size exclusion chromatography. The protein pattern suggests that SEC purification and enrichment of RNase P activity of HIC<sub>0.2</sub> samples was more efficient than SEC purification of HIC<sub>0.4</sub> samples, considering that the lower protein band (15 – 25 kDa) in the SEC<sub>0.2</sub> sample (arrow) was later identified as Aq\_880 by Western Blot analysis. AEC: anion exchange chromatography; HIC: hydrophobic interaction chromatography; SEC: size exclusion chromatography.

### Quantitative mass spectrometry (liquid samples with RNase P activity)

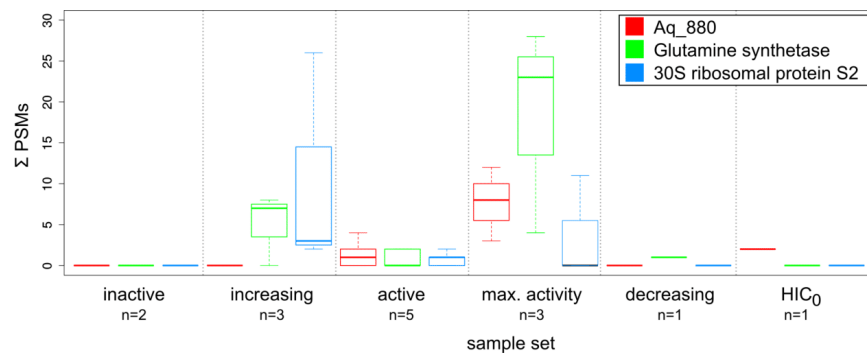
According to the results of the activity assays from RNase P-enriched SEC<sub>0.2</sub> and SEC<sub>0.4</sub> fractions, the elution fractions were classified as "-/+" (very low RNase P activity, increasing), "+" (increasing or decreasing activity), "++" (maximum activity) "+/-" (very low activity, decreasing). Fractions from each category were analyzed applying label-free quantitative mass spectrometry.



**Figure S9.** SEC fractions were tested for RNase P activity. Reaction products were analyzed by denaturing 20% PAGE in the order of elution from the Superose 6 column (here illustrated for SEC<sub>0.4</sub> fractions). "-/+": very low RNase P activity, increasing; "+": increasing or decreasing activity; "++": maximum activity; "+/-": very low RNase P activity, decreasing. P: positive control, cleavage of pre-tRNA<sup>Gly</sup> by *E. coli* P RNA as in Fig. S2; N: negative control (ddH<sub>2</sub>O instead of enzyme).

The data obtained from quantitative mass spectrometry analysis were compared to the RNase P activity of the analyzed fractions. The *Peptide Spectrum Matches* value ( $\sum$ PSM) was used as an indication for the quantitative amount of a certain protein in the analyzed fraction. The most abundant proteins in fractions containing maximum RNase P activity were (I) a hypothetical protein encoded by gene *aq\_880*, (II) glutamine synthetase and (III) 30S ribosomal protein (r-protein) S2. Included in the analysis was a sample that was not SEC-purified following HIC purification. This sample contained the pooled and concentrated fractions that eluted during HIC in step 5 at 0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (= HIC<sub>0</sub>). After concentration, RNase P activity was also detectable in these fractions. The sample HIC<sub>0</sub> was analyzed by mass spectrometry to search for common proteins also detectable in the SEC<sub>0.2</sub> and SEC<sub>0.4</sub> fractions displaying maximum RNase P activity (Fig. S10).





**Figure S10.** Boxplots for the three most abundant proteins in the fractions containing high RNase P activity (glutamine synthetase, Aq\_880 and r-protein S2). The sum of *Peptide Spectrum Matches* (PSMs) for the specific protein is plotted over the height of RNase P activity of the analyzed SEC fraction according to Fig. S9 (inactive; -/+ : low activity, increasing; + : active; ++ : maximum (max.) activity; +/- : low activity, decreasing). Inactive: based on one SEC<sub>0,2</sub> and one SEC<sub>0,4</sub> sample (n = 2); increasing: based on two SEC<sub>0,2</sub> and one SEC<sub>0,4</sub> sample (n = 3); active: based on three SEC<sub>0,2</sub> and two SEC<sub>0,4</sub> samples (n = 5); max. activity: based on two SEC<sub>0,2</sub> and one SEC<sub>0,4</sub> sample (n = 3); decreasing: based on a single SEC<sub>0,4</sub> sample (n = 1); HIC<sub>0</sub> represented a single pool of concentrated fractions (with weak RNase P activity) that eluted in step 5 during HIC. Whiskers indicate the most extreme PSM sums found, boxes indicate the average range of 50% of all respective PSM sums in which the median of all values is shown as a thick line. For example, in the max. activity fraction, Aq\_880 was found in three independent samples with PSM sums of 3, 8 and 12. Thus, the most extreme values are 3 and 12 (whiskers), while the average of all values (box) is in the range of 5.5 to 10 (means of 3 and 8, and 8 and 12) with a median of 8 (thick line).

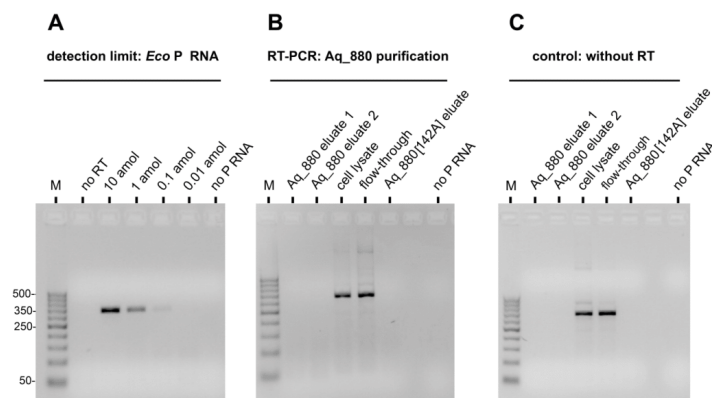
As a result, only the protein Aq\_880 was identified in both, the SEC fractions containing the highest RNase P activity and the concentrated HIC<sub>0</sub> fraction with moderate RNase P activity.

Apart from Aq\_880, other abundant proteins were also identified in the SEC fractions containing maximum RNase P activity (in order of decreasing frequency): glutamine synthetase, the 30S ribosomal protein S2, the polynucleotide phosphorylase (PNPase), the N utilization substance protein B homolog (NusB) and Aq\_707. Other ribosomal proteins were detected in fractions with increasing RNase P activity, but not in fractions with maximum activity. PNPase, NusB and Aq\_707 were less abundant than Aq\_880 and S2.

### Excluding an *E. coli* RNase P RNA contamination in recombinant Aq\_880 preparations

#### RT-PCR

RT-PCR analysis excluded a contamination with *E. coli* RNase P as the source of pre-tRNA processing by recombinant Aq\_880 preparations (Fig. S11).

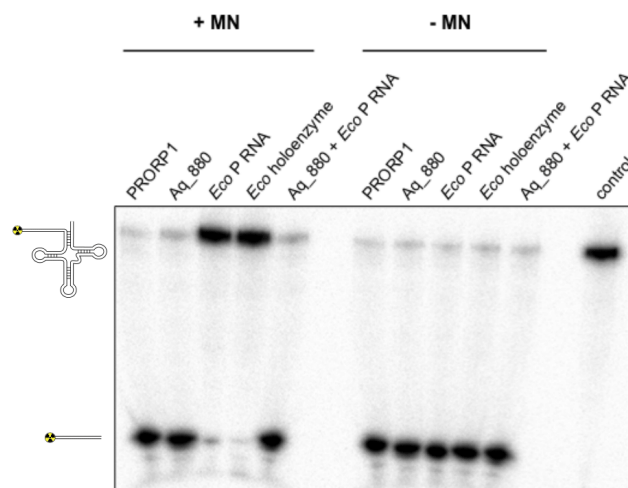


**Figure S11:** Detection of *E. coli* RNase P RNA (P RNA) by RT-PCR. All samples were extracted with phenol/chloroform and precipitated with ethanol before (RT-)PCR analysis. Three  $\mu$ l from each reaction were separated on a 2% agarose gel and stained with ethidium bromide. **(A)** First, the detection limit of 0.1 amol ( $10^{-19}$  mol) P RNA was determined using a concentration range of 10 amol ( $10^{-17}$  mol) to 10 zmol ( $10^{-20}$  mol) of *in vitro*-transcribed *E. coli* P RNA. As a control, the RT-PCR reaction was performed in the absence of reverse transcriptase (no RT) or P RNA template (no P RNA) to exclude DNA contamination or unspecific amplification. M, 50-bp DNA ladder (Roth, Germany). **(B)** Ni-NTA-purified fractions of recombinant Aq\_880 and Aq\_880[D142A] were analyzed for the presence of *E. coli* P RNA by RT-PCR; 30 ng (1.25 pmol) batch-purified Aq\_880 (Aq\_880 eluate 1), FPLC-purified Aq\_880 (Aq\_880 eluate 2) or corresponding Aq\_880[D142A] fractions were tested with primers specific for *E. coli* P RNA. As positive controls, 2.5  $\mu$ l cell lysate (before loading onto the column) and 2.5  $\mu$ l of the flow-through fraction were included. **(C)** In parallel, all fractions were tested for DNA contaminations in corresponding reactions but omitting reverse transcriptase. Bands in the cell lysate and flow-through indicate that DNA in these samples is also responsible for the signals in panel B.

#### Treatment with micrococcal nuclease (MN)

The conclusion that *A. aeolicus* RNase P Aq\_880 is a protein-only enzymatic activity is corroborated by *in vitro* experiments with micrococcal nuclease (MN). The treatment with MN had no effect on the pre-tRNA processing activity of recombinant Aq\_880. The same result

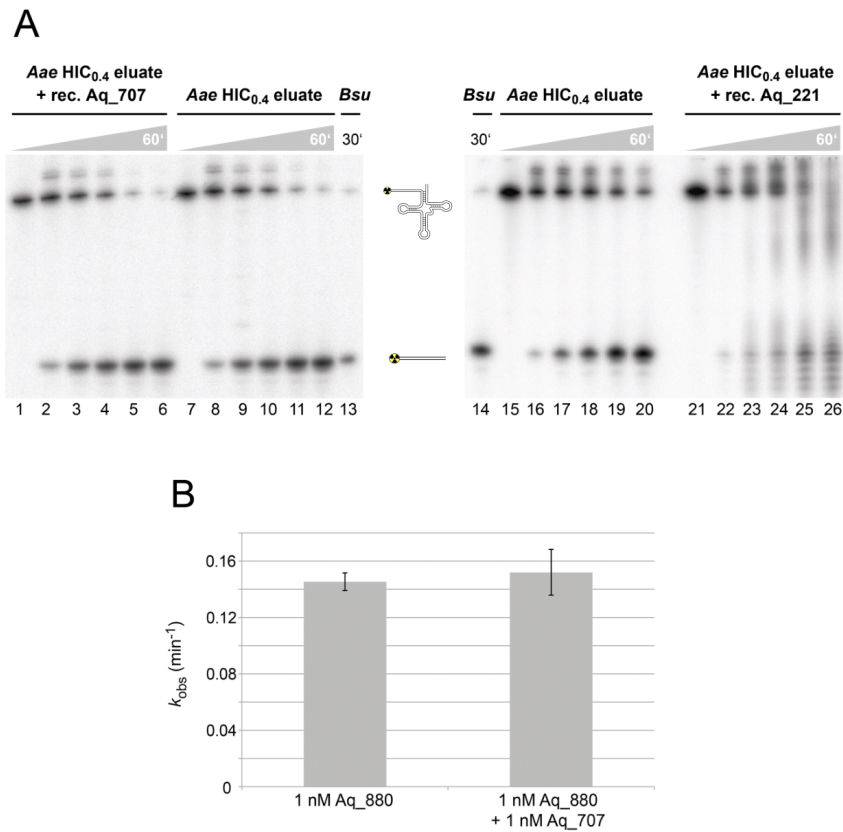
was seen in a parallel control experiment with PRORP1, the organellar protein-only RNase P from *Arabidopsis thaliana*. In contrast, incubation of *E. coli* P RNA or RNase P holoenzyme with MN completely abolished the pre-tRNA processing activity of this RNA-based enzyme. Unspecific binding of MN to pre-tRNA could make the cleavage site inaccessible for RNase P ("substrate masking", (46). Thus, a carrier RNA (*Bacillus subtilis* 6S-1 RNA) was added in 500-fold molar excess over pre-tRNA to processing assays to prevent unspecific binding of MN to pre-tRNA.



**Figure S12.** Processing of pre-tRNA<sup>Gly</sup> by RNase P enzymes after treatment with micrococcal nuclease (MN). PRORP1: *A. thaliana* PRORP1; Aq\_880: *A. aeolicus* Aq\_880; Eco P RNA: *E. coli* RNase P RNA in the presence of 100 mM Mg<sup>2+</sup>; Eco holoenzyme: *E. coli* RNase P holoenzyme in presence of 10 mM Mg<sup>2+</sup>; Aq\_880 + Eco P RNA: *A. aeolicus* Aq\_880 plus *E. coli* P RNA; control: incubation of substrate in the absence of enzyme. + MN: RNase P samples pretreated with MN before pre-tRNA<sup>Gly</sup> processing assays; - MN: as samples "+ MN", but omission of MN. For experimental details, see Supplementary Materials and Methods below.

#### Addition of recombinant Aq\_707 / Aq\_221 (PNPase) to *A. aeolicus* fractions with enriched RNase P activity

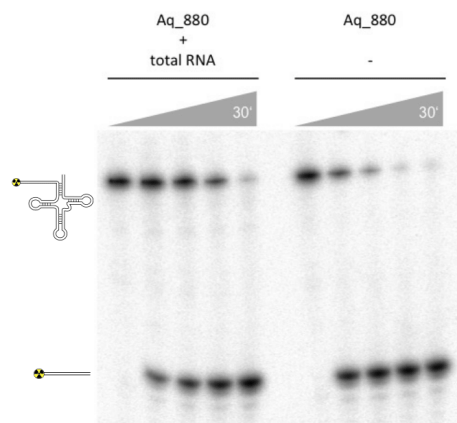
Since Aq\_707 and Aq\_221 (PNPase) copurified with RNase P activity and Aq\_707 formed stable complexes with Aq\_880 in SDS gels (band #6 in Fig. 1B of the main manuscript), we also tested if these recombinantly produced *A. aeolicus* proteins support RNase P activity. However, neither Aq\_707 nor PNPase stimulated the RNase P activity of Aq\_880. Not unexpected, PNPase caused some unspecific 3'-exonucleolytic degradation of the pre-tRNA and the 5'-leader cleavage product.



**Figure S13. (A)** Effect of recombinant *A. aeolicus* proteins Aq\_707 or Aq\_221 (PNPase) on RNase P activity of partially purified RNase P (HIC<sub>0.4</sub>) fractions from *A. aeolicus*. Aliquots were withdrawn after 1, 5, 10, 30 and 60 min. Lanes 1, 7, 15 and 21: pre-tRNA<sup>Gly</sup> incubated for 60 min in reaction buffer lacking the HIC<sub>0.4</sub> eluate and recombinant proteins was used as negative control (lanes 1, 7, 15, 21). Bsu: pre-tRNA<sup>Gly</sup> incubated for 30 min in the presence of 10 nM *B. subtilis* RNase P holoenzyme. For experimental details, see Suppl. Materials and Methods, "Cell lysate supplemented with recombinant proteins", p.37. **(B)** Cleavage of 5 nM pre-tRNA<sup>Gly</sup> by 1 nM recombinant Aq\_880, either in the absence or presence of 1 nM Aq\_707 in buffer F at 37°C;  $k_{obs}$  values are mean values derived from three independent experiments (single exponential decay based on time points at 0, 5, 15, 35 and 55 min), error bars are standard deviations of the mean. For further details, see Suppl. Materials and Methods, paragraph "Activity assay with recombinant Aq\_880 variants", p. 35.

### Activity of recombinant Aq\_880 in the presence of *A. aeolicus* total RNA

The preceding experiments provided no evidence for *A. aeolicus* RNase P activity depending on an essential RNA component. However, the possibility remained that RNA component(s) might support Aq\_880 activity by exerting auxiliary structural or regulatory functions. Therefore, we analyzed Aq\_880 activity under single-turnover conditions in the presence and absence of total *A. aeolicus* RNA either prepared as reported (8) (Method 1 without lysozyme) or as described in Suppl. Materials and Methods ("A. *aeolicus* total RNA preparation", acetic acid plus phenol/chloroform extraction). Such experiments (Fig. S14) showed that addition of *A. aeolicus* total RNA even had an inhibitory effect on the pre-tRNA processing activity of Aq\_880. The relative rate constant  $k_{\text{rel}}$  is determined as the ratio of rate constants for Aq\_880 with and without supplementation of total RNA ( $k_{\text{rel}} = k_{\text{obs}(+\text{RNA})} / k_{\text{obs}(-\text{RNA})}$ );  $k_{\text{rel}}$  was 0.47 at 37°C and 0.23 at 70°C assay temperature (mean value of two independent experiments each).



**Figure S14.** RNase P activity of Aq\_880 (50 nM) in the presence and absence of *A. aeolicus* total RNA (20 ng/μl; prepared by acetic acid plus phenol/chloroform extraction) under single-turnover conditions at 37°C. Very similar results were obtained when processing assays were performed at 70°C or with total *A. aeolicus* RNA prepared according to (8). For experimental details, see Suppl. Materials and Methods, paragraph "Aq\_880 supplemented with total RNA", p. 37.

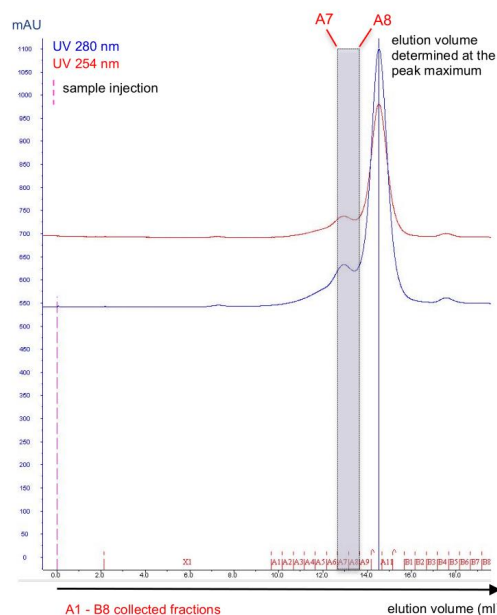
### Thermostability of recombinant Aq\_880

As the hyperthermophile *A. aeolicus* thrives at temperatures of up to 95°C, the thermostability of recombinant Aq\_880 was determined in a cleavage assay using *in vitro*-transcribed pre-tRNA<sup>Gly</sup> from the thermophilic bacterium *T. thermophilus*. In such assays, processing by recombinant Aq\_880 was still observable at 70°C, but only residual activity was seen at 75°C and no activity at higher temperatures. The rate constant at 70°C ( $k_{\text{obs}} = 0.064 \text{ min}^{-1}$ ) was 15-fold lower than at 37°C ( $k_{\text{obs}} = 0.90 \text{ min}^{-1}$ ). As the rate decrease may have included contributions from increased dissociation rates of enzyme-substrate complexes and from denaturation of the pre-tRNA *in vitro* transcript lacking post-transcriptional nucleoside modifications that were shown to confer increased thermostability (37), we pursued a setup in which Aq\_880 was preincubated for 10 min at 85°C before performing the pre-tRNA processing assay at 37°C. The  $k_{\text{obs}}$  slightly decreased from 1.17 min<sup>-1</sup> (no preincubation at 85°C) to 0.79 min<sup>-1</sup> with preincubation at 85°C (see Fig. 2F of the main manuscript). Thus, the preincubation step decreased the enzymatic activity by not more than ~30%, consistent with thermostability of Aq\_880. Applying the same heat treatment to the *E. coli* RNase P holoenzyme essentially resulted in a complete loss of its enzymatic activity ( $k_{\text{obs}} = 0.005 \text{ min}^{-1}$ ) compared to preincubation at 37°C ( $k_{\text{obs}} = 1.37 \text{ min}^{-1}$ ). The same kind of analysis was performed with a native HIC<sub>0.4</sub> fraction (see Fig. S3), giving very similar results (see Fig. 2F of the main manuscript).

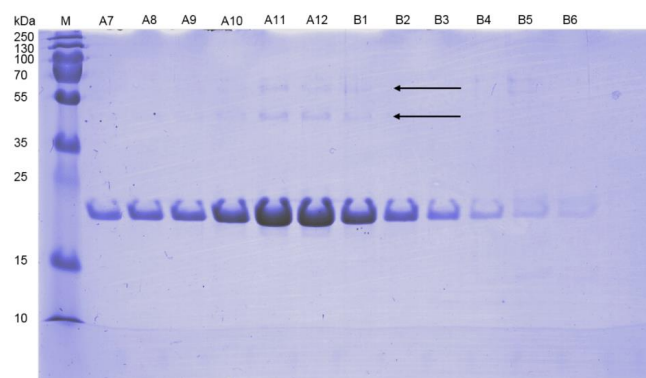
### Oligomerization of recombinant Aq\_880

The elution volume of recombinant Aq\_880 on a calibrated Superose 6 column corresponded to a mass of approximately 420 kDa (Fig. S15). The protein eluted closely after apoferritin (443 kDa) that was used for column calibration. The result agrees with the estimates of about 350 kDa based on the elution volume of partially purified SEC fractions with maximum RNase P activity during the third purification step (Superose 6 column, Fig. S5 to S7).

SDS-PAGE analysis of the fractions eluting from the Superose 6 column loaded with recombinant Aq\_880 (Fig. S15) revealed weak protein bands at about 50 and 70 kDa in the fractions with the highest concentration of Aq\_880 (Fig. S16, arrows). In the minor band at around 70 kDa, exclusively *A. aeolicus* Aq\_880 was identified by mass spectrometry. This correlates with the size of the extra band in the Aq\_880 Western blot (Fig. 3B of the main manuscript; migrating there somewhat faster because of lack of the His tag). The band migrating at ~ 50 kDa was identified as the *E. coli* maltoporin precursor. This indicates that Aq\_880 forms stable trimers that are able to persist during SDS-PAGE. Considering that the major elution peak of Aq\_880 corresponded to a mass of approximately 420 kDa, it is possible that Aq\_880 occurs in oligomeric forms consisting of 3 hexamers or 6 trimers.



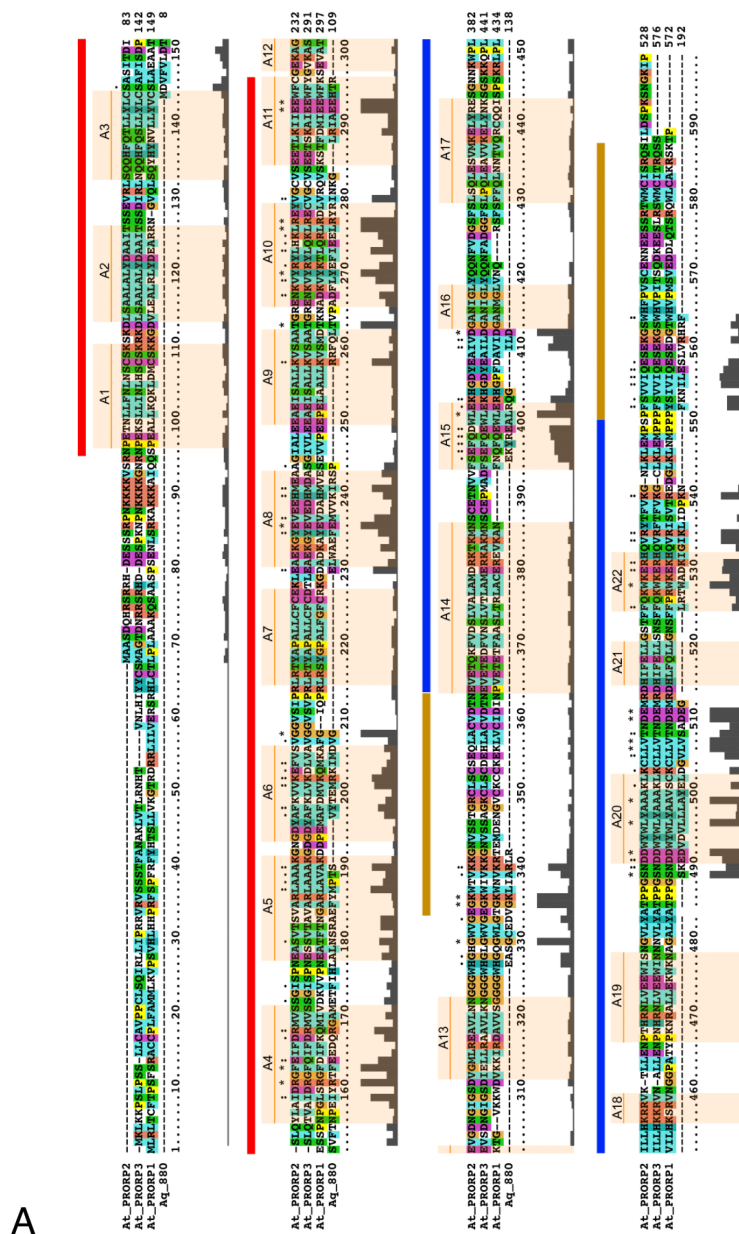
**Figure S15.** Elution profile of recombinant Aq\_880 in buffer S (50 mM Tris-HCl pH 7.45, 150 mM NaCl, 4.5 mM Mg(OAc)<sub>2</sub>, 5 mM DTT<sub>2</sub>) on a Superose 6 column (flow rate: 0.3 ml/min). The elution volume was determined at the peak maximum. Fractions of 0.5 ml were collected and analyzed by SDS-PAGE. The fractions collected from the small peak preceding the main peak were analyzed as well (fractions A7 and A8).



**Figure S16.** SDS-PAGE analysis of fractions A7 – B6 eluting from the Superose 6 column loaded with recombinant C-terminally His-tagged Aq\_880 (FPLC-purified). Protein bands were detected by sensitive colloidal Coomassie staining. The corresponding chromatogram is shown in Fig. S15. The arrows point to the additional weak signals at about 50 and 70 kDa that became visible in the peak fractions A10-B1. Only a single band of the size of monomeric Aq\_880 (23.7 kDa, with C-terminal His tag) was detectable in fractions A7 and A8 that were collected from the small peak preceding the main peak. 15% SDS-PAGE; M: Protein marker.

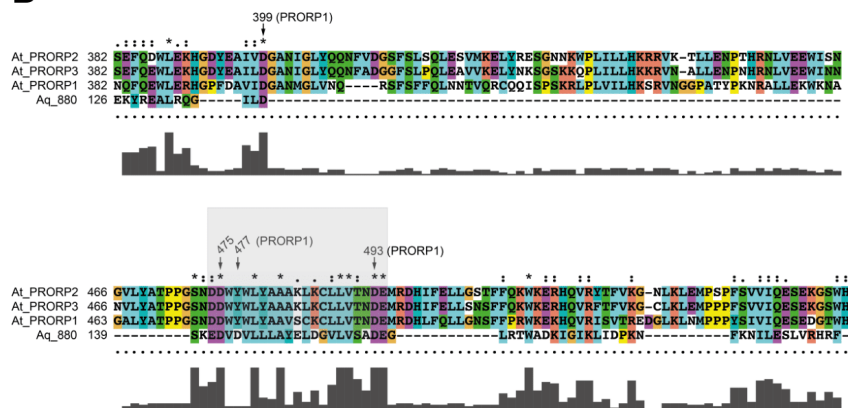


## Identification of active site residues in Aq\_880

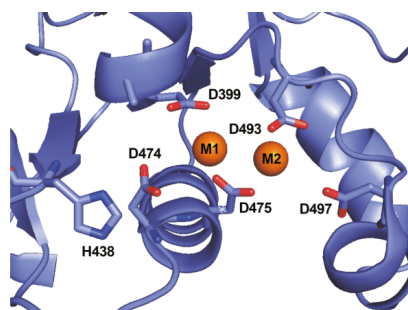




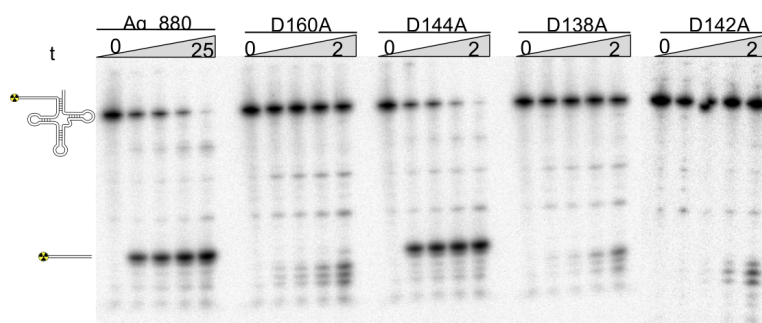
B



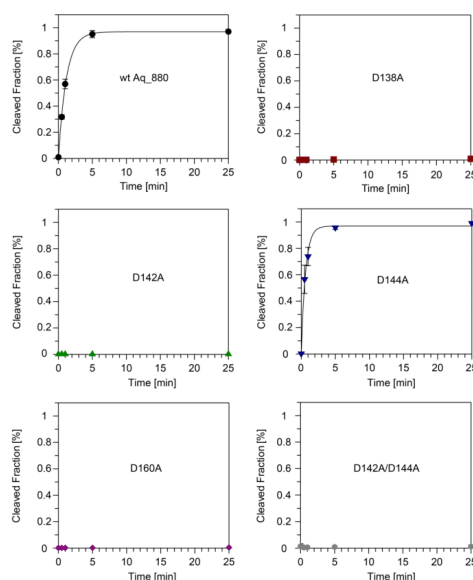
**Figure S17. (A)** Clustal Omega (38) alignment of full-length *A. thaliana* PRORP1-3 (At\_PRORP1-3) and Aq\_880. The  $\alpha$ -helical elements inferred from the X-ray structure of At\_PRORP1 (21) are numbered A1–A22 above the sequences. The amino acid positions of the pentatricopeptide repeat domain (red boxes at the top, residues 95–292 of At\_PRORP1), the central domain (golden boxes, residues 328–357 and 534–570 of At\_PRORP1) and the metallonuclease domain (blue boxes at the top, residues 358–533 of At\_PRORP1) are indicated. **(B)** Section of the alignment that revealed similarities of Aq\_880 to *A. thaliana* PRORP1-3 in their metallonuclease domains (Nedd4-BP1, YacP nucleases = NYN domains); in the area highlighted in gray, candidates for catalytic aspartates of Aq\_880 were identified. In addition, D138 of Aq\_880 was considered as a catalytic aspartate corresponding to D399 of At\_PRORP1.



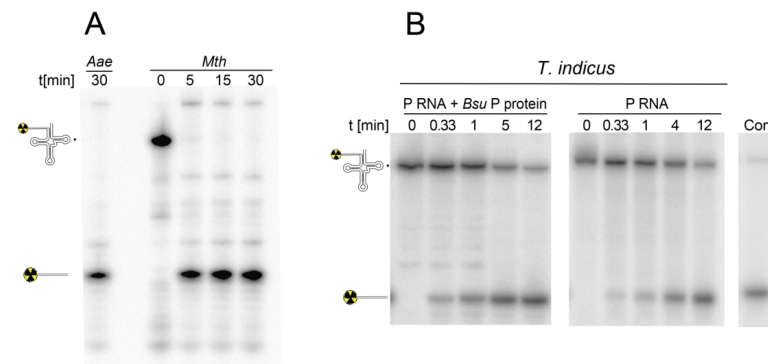
**Figure S18.** Illustration of the catalytic metal ion binding sites in *A. thaliana* PRORP1. Based on the alignment of PRORP enzymes and Aq\_880 (Fig. S17), Aq\_880 residues D138, D142 and D160 were inferred to correspond to D399, D475 and D493 of PRORP1. The aspartic acid residue chains in the active center of PRORP1 are shown as sticks. The side chains of D399, D474, D475 and D493 are supposed to coordinate two catalytic metal ions (orange spheres). The illustration is based on the *A. thaliana* PRORP1 crystal structure (21).



**Figure S19.** *In vitro* pre-tRNA processing assay under single-turnover conditions (50 nM enzyme, < 1nM substrate) to investigate the function of Aq\_880 variants with mutation of aspartic acid residues (to alanine) potentially involved in catalysis. All proteins carried a C-terminal His tag; assays were performed at 37°C. The pre-tRNA<sup>Gly</sup> (upper band) was 5'-[<sup>32</sup>P]-end-labeled. Mutants D160A, D138A and D142A were inactive, only mutant D144A was similarly active as wild-type Aq\_880.



**Figure S20.** Graphical illustration of *in vitro* pre-tRNA<sup>Gly</sup> processing assays of the type presented in Fig. S19. The fraction of processed pre-tRNA was plotted against the time [min]. The values are mean values of three independent single-turnover experiments performed at 37°C. Essentially no activity was detectable for the mutants D138A, D142A, D160A and D142A/D144A. For wild-type Aq\_880 and the mutant D144A enzyme, we calculated  $k_{\text{obs}}$  values of  $0.93 \pm 0.05 \text{ min}^{-1}$  and  $1.6 \pm 0.13 \text{ min}^{-1}$  under these test conditions.

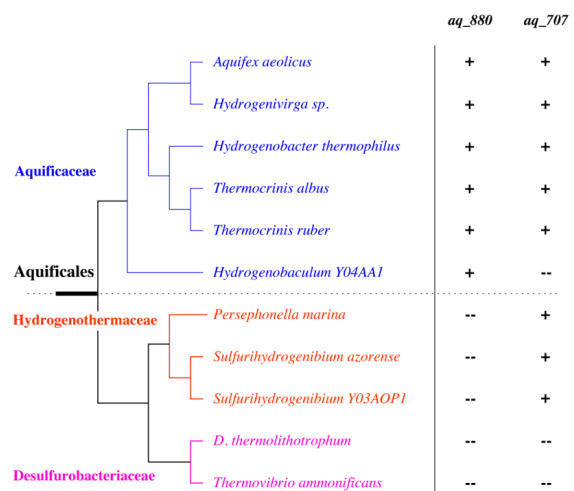


**Figure S21.** Activity assays for RNA-based and HARP RNase P from *Thermodesulfatator indicus* (bacterium) and HARP from *Methanothermobacter thermautotrophicus* Delta H (archaeon). **(A)** Enzymatic activity of archaeal HARP from *M. thermautotrophicus* (Mth) was analyzed as described in the Suppl. Materials and Methods, paragraph "Activity assays using other HARP proteins" (p. 38). Samples were separated by 20% denaturing PAGE. A corresponding assay with Aq\_880 (Aae, 30 min of incubation) served as control (see Suppl. Materials and Methods, paragraph "Activity assay with recombinant Aq\_880 variants", p. 35). **(B)** Enzymatic activity of the chimeric holoenzyme composed of *T. indicus* P RNA and *B. subtilis* P protein and *T. indicus* P RNA alone was analyzed as described in Suppl. Materials and Methods, paragraph, "Activity assays with other bacterial RNase P RNAs" (p. 38). As control (lane Con), an RNase P holoenzyme consisting of *E. coli* P RNA and *B. subtilis* P protein was reconstituted as described for the chimeric *T. indicus* enzyme and incubated with the substrate for 12 min before gel loading. RNA-based RNase P of *M. thermautotrophicus* was not tested *in vitro* as the enzyme was previously characterized in much detail (3) (15) (16).

## Bioinformatic analysis

### Evaluation of proteins identified in quantitative and qualitative mass spectrometry analyses

Of all proteins identified in the quantitative and qualitative mass spectrometry analyses, the hypothetical protein Aq\_880 appeared as the most promising one, because it was the most abundant candidate and it was the only protein that was detected in both, SEC<sub>0.2</sub>/SEC<sub>0.4</sub> fractions with maximum RNase P activity and in the minor HIC<sub>0</sub> fractions with weak RNase P activity (Fig. S10). The structure prediction of Aq\_880 revealed similarities to a solution structure of an endonuclease and showed similarity to a PIN (PiIT N-terminal) domain. Since PIN domains function as ribonucleases (2) and the catalytic metallonuclease domain of the protein-only RNase P (PRORP) belongs to the PIN domain-like fold superfamily (44), the Aq\_880 protein was considered to function as an RNase P in *A. aeolicus*. The importance of the hypothetical protein Aq\_880 is underlined by its exclusive appearance in the *Aquificaceae* branch lacking any bacterial-like RNase P (25) (27) (Fig. S22).



**Figure S22.** Identification of *aq\_880* and *aq\_707* homologs within the *Aquificales*. The search was performed with tblastn applying an *e* value threshold of  $10^{-5}$ . The phylogenetic tree was adapted from (25). *Aq\_707* was included in the analysis because (i) the protein copurified with RNase P activity, (ii) *Aq\_707* formed stable complexes with *Aq\_880* (see main manuscript), and (iii) the protein was predicted as a tRNA methyltransferase homolog (see text below).

### Analysis of structural domains

The most abundant proteins in the SEC fractions with maximum RNase P activity were further analyzed by structural bioinformatics using PHYRE2 in the intensive mode (22). As we were searching for *A. aeolicus* RNase P, the focus was put on RNA-binding properties and endonucleolytic function. The alignment and structure prediction of *Aq\_880* showed similarities to the PIN (PilT N-terminal) domain that belongs to the PIN domain-like superfamily. The metalloendonuclease domain of the protein-only RNase P (PRORP1) contains a NYN (N4BP1, YacP-like Nuclease) domain (1) (21) which also belongs to the PIN domain-like superfamily. The identity value between *Aq\_880* and the PIN domain was only 28% but the confidence of the result was 97.5%. The relatively low identity might be attributable to only part of the PIN domain protein serving as template for the alignment with *Aq\_880* that essentially is a PIN domain only. The alignment resulted in additional hits for RNA binding (functionally unknown protein from *Methanocaldococcus jannaschii*, 21% identity, 98.5% confidence) and metal binding (endonuclease Nob1 from *Pyrococcus horikoshii*, 33% identity, 98.9% confidence) proteins. Regarding the fact that all RNase P enzymes identified up to date are metalloendonucleases, *Aq\_880* became a top candidate for an *A. aeolicus* RNase P component. The analysis of the *A. aeolicus* PNPase pointed to a structural similarity between the PNPase and bacterial P proteins. (41) already analyzed the structure

of the PNPase from *Streptomyces antibioticus* and described bacterial RNase P proteins as likely structural homologs of PNPase/GPSI domains. According to the SCOP (29) and the SUPERFAMILY database (12), PNPase/GPSI domains are a member of the *Ribonuclease PH domain 1-like* family, which, together with the *RNase P protein* family, belong to the *Ribosomal protein S5 domain 2-like* superfamily. As PNPase is known to be an exonuclease we followed this track with regard to structural similarity. The analysis of the hypothetical protein Aq\_707 revealed partial similarities (20 – 22%, confidence: 100%) with MnmC proteins from *E. coli* and *Yersinia pestis* that are involved in tRNA 5-methylaminomethyl-2-thiouridine biosynthesis. In *E. coli*, MnmC is a bifunctional tRNA modification enzyme with methyltransferase activity. Since one subunit (TRMT10C) of the human mitochondrial RNase P is a methyltransferase (20), (44), the hypothetical protein Aq\_707 moved into our focus despite the fact that the methyltransferase homolog has already been identified in *A. aeolicus* (23). Nonetheless, Aq\_707 remained in our focus based on its potential to be a protein that acts on tRNAs.

Furthermore, we analyzed the glutamine synthetase. Although the structure and function of this enzyme is known, we performed structure predictions and alignments, taking into consideration that *A. aeolicus* glutamine synthetase might have acquired additional, yet unknown, functions during evolution. However, the results neither provided support in favor of such a possibility nor did we find any evidence for an RNA binding domain. Regarding NusB, which plays an important role in transcriptional antitermination, its RNA-binding properties and sites are known (9) (39). Its structure analysis failed to give any hints on additional functions that might be related to RNase P activity.

#### **Occurrence of protein candidates within the *Aquificales* and alignment of bacterial and archaeal Aq\_880 homologs**

Despite intensive bioinformatic efforts neither a gene for a bacterial-type RNase P protein subunit nor an RNA subunit could be identified in *A. aeolicus* or any other member of the

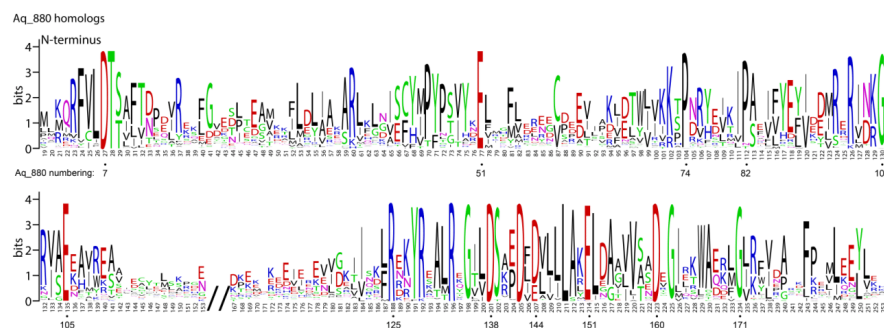




**Figure S23.** Clustal Omega multiple sequence alignment of Aq\_880 homologs among Bacteria and Archaea. For details, see Supplementary Materials and Methods.

family of the *Aquificaceae*. However, bacterial-type RNase P enzymes have been identified in other *Aquificales*, including all species that belong to the families *Desulfurobacteriaceae* and *Hydrogenothermaceae* which, together with the *Aquificaceae*, form the order *Aquificales* (25) (26). We thus assumed that the RNase P from *Aquificaceae* differs from the RNase P of *Desulfurobacteriaceae* and *Hydrogenothermaceae*. Therefore, the proteins identified in RNase P-enriched *A. aeolicus* fractions were analyzed for commonalities among *Aquificaceae* and concomitant differences to homologs from the *Desulfurobacteriaceae* and *Hydrogenothermaceae*. Yet, most informative was the presence or absence of proteins in the three families belonging to the *Aquificales* (Fig. S22). The glutamine synthetase, NusB, PNPase and ribosomal protein S2 were found to be represented in all three families, which was expected for housekeeping genes. The hypothetical protein encoded by gene *aq\_707* was not detected in the *Desulfurobacteriaceae* and in the species *Hydrogenobaculum* sp. Y04AAS1 (Fig. S22), which belongs to the *Aquificaceae*, but was identifiable in the *Hydrogenothermaceae* expressing canonical bacterial RNase P enzymes (26). In contrast, as mentioned above, *aq\_880* homologs were exclusively found in all analyzed members of the *Aquificaceae*, but not in any of the two other families of the *Aquificales*.

Aq\_880 homologs were identified in many Archaea and a small group of bacteria beyond the *Aquificaceae* (Fig. S23, S24 and Table S1). For more information, see the main manuscript.



**Figure S24.** WebLogo to illustrate amino acid conservation among the Aq\_880 homologs aligned in Fig. S23. In addition to the numbering based on the alignment of Fig. S23, residue numbers in Aq\_880 are indicated for some of the conserved amino acids. For more details, see Suppl. Materials and Methods.

Species	HARP	RNA	RnpA	Rpp29	Pop5	Rpp30
<b>Bacteria</b>						
<b>Aquificae -&gt; Aquificae -&gt; Aquificales -&gt; Aquificaceae</b>						
<i>Aquifex aeolicus</i> VF5	yes				yes	yes
<i>Hydrogenobacter thermophilus</i> TK 6	yes				yes	yes
<i>Hydrogenobaculum</i> HO	yes				yes	yes
<i>Hydrogenobaculum</i> SN	yes				yes	yes
<i>Hydrogenobaculum</i> Y04AAS1	yes				yes	yes
<i>Thermocrinis albus</i> DSM 14484	yes				yes	yes
<b>Proteobacteria -&gt; Gammaproteobacteria -&gt; Chromatiales -&gt; Chromatiaceae</b>						
<i>Nitrosococcus halophilus</i> Nc4	yes	yes	yes		(yes)	yes
<i>Nitrosococcus oceani</i> ATCC 19707	yes	yes	yes		(yes)	
<b>Proteobacteria -&gt; Gammaproteobacteria -&gt; Chromatiales -&gt; Ectothiorhodospiraceae</b>						
<i>Alkalilimnicola ehrlichii</i> MLHE 1	yes	yes	yes		(yes)	
<i>Halorhodospira halophila</i> SL1	yes	yes	yes			
<i>Thioalkalivibrio</i> K90mix	yes	yes	yes		yes	
<i>Thioalkalivibrio nitratireducens</i> DSM 14787	yes	yes	yes		yes	
<b>Proteobacteria -&gt; Gammaproteobacteria -&gt; Methylococcales</b>						
<i>Methylohalobius crimeensis</i>	yes	yes	yes			yes
<b>Verrucomicrobia -&gt; Methylacidiphilales -&gt; Methylacidiphilaceae</b>						
<i>Methylacidiphilum inferorum</i> V4	yes	yes	yes		(yes)	yes
<b>Thermodesulfobacteria -&gt; Thermodesulfobacteria -&gt; Thermodesulfobacteriales -&gt; Thermodesulfobacteriaceae</b>						
<i>Thermodesulfatator indicus</i> DSM 15286	yes	yes	yes			yes
<i>Thermodesulfobacterium geofontis</i> OPF15						
(initially termed <i>Thermodesulfobacterium</i> OPB45)	yes	yes	yes			yes
<b>Nitrospirae -&gt; Nitrospira -&gt; Nitrospirales -&gt; Nitrospiraceae</b>						
<i>Thermodesulfobacterium yellowstonii</i> DSM 11347	yes					yes
<b>Archaea</b>						
<b>Thermococci -&gt; Thermococcales -&gt; Thermococcaceae</b>						
<i>Pyrococcus</i> NA2	yes	yes		yes	yes	yes
<i>Pyrococcus</i> ST04	yes	yes		(yes)	yes	yes
<i>Pyrococcus abyssi</i> GE5	yes	yes		yes	yes	yes
<i>Pyrococcus furiosus</i> COM1	yes	yes		yes	yes	yes
<i>Pyrococcus furiosus</i> DSM 3638	yes	yes		yes	yes	yes
<i>Pyrococcus horikoshii</i> OT3	yes	yes		yes	yes	yes
<i>Pyrococcus yamanosii</i> CH1	yes	yes		yes	yes	yes
<i>Thermococcus</i> 4557	yes	yes		yes	yes	yes
<i>Thermococcus</i> AM4	yes	yes		yes	yes	yes
<i>Thermococcus</i> CL1	yes	yes		(yes)	yes	yes



<i>Thermococcus barophilus</i> MP	yes	yes	yes	yes	yes
<i>Thermococcus gammatolerans</i> EJ3	yes	yes	yes	yes	yes
<i>Thermococcus kodakarensis</i> KOD1	yes	yes	yes	yes	yes
<i>Thermococcus litoralis</i> DSM 5473	yes	yes	yes	yes	yes
<i>Thermococcus onnurineus</i> NA1	yes	yes	yes	yes	yes
<i>Thermococcus sibiricus</i> MM 739	yes	yes	yes	yes	yes
<b>Halobacteria -&gt; Halobacteriales -&gt; Halobacteriaceae</b>					
<i>Haloarcula marismortui</i> ATCC 43049	yes	yes	(yes)	yes	yes
<i>Halorhabdus tiamatea</i> SARL4B	yes	yes	yes	yes	yes
<i>Halorhabdus utahensis</i> DSM 12940	yes	yes	yes	yes	yes
<b>Halobacteria -&gt; Haloferacales -&gt; Haloferacaceae</b>					
<i>Haloferax mediterranei</i> ATCC 33500	yes	yes	yes	yes	yes
<i>Haloferax volcanii</i> DS2	yes	yes	yes	yes	yes
<i>Halogeometricum borinquense</i> DSM 11551	yes*	yes	(yes)	(yes)*	(yes)
<b>Halobacteria -&gt; Natrilbales -&gt; Natrilbaceae</b>					
<i>Halopiger xanaduensis</i> SH 6	yes	yes	yes	yes	yes
<i>Haloterrigena turkmenica</i> DSM 5511	yes	yes	yes	yes	yes
<i>Natrialba magadii</i> ATCC 43099	yes	yes	yes	yes	yes
<i>Natrinema J7</i>	yes	yes	yes	yes	yes
<i>Natrinema pellirubrum</i> DSM 15624	yes	yes	yes	yes	yes
<i>Natronococcus occultus</i> SP4	yes	yes	yes	yes	yes
<b>Archaeoglobi -&gt; Archaeoglobales -&gt; Archaeoglobaceae</b>					
<i>Archaeoglobus fulgidus</i> DSM 4304	yes	yes	yes	yes	yes
<i>Archaeoglobus profundus</i> DSM 5631	yes	yes	yes	yes	yes
<i>Archaeoglobus sulfaticallidus</i> PM70 1	yes	yes	yes	yes	yes
<i>Archaeoglobus veneficus</i> SNP6	yes	yes	yes	yes	yes
<i>Ferroglobus placidus</i> DSM 10642	yes	yes	yes	yes	yes
<b>Methanobacteria -&gt; Methanobacteriales -&gt; Methanobacteriaceae</b>					
<i>Methanobacterium</i> AL 21	yes	yes	yes	yes	yes
<i>Methanobacterium</i> MB1	yes	yes	yes	yes	yes
<i>Methanobacterium</i> SWAN 1	yes	yes	yes	yes	yes
<i>Methanothermobacter marburgensis</i> Marburg	yes	yes	yes	yes	yes
<i>Methanothermobacter thermautotrophicus</i> Delta H	yes	yes	yes	yes	yes
<b>Methanobacteria -&gt; Methanobacteriales -&gt; Methanothermaceae</b>					
<i>Methanothermus fervidus</i> DSM 2088	yes	yes	yes	yes	yes
<b>Methanococci -&gt; Methanococcales -&gt; Methanocaldococcaceae</b>					
<i>Methanocaldococcus</i> FS406 22	yes	yes	yes	yes	yes
<i>Methanocaldococcus fervens</i> AG86	yes	yes	yes	yes	yes

<i>Methanocaldococcus infernus</i> ME	yes	yes	yes	yes	yes
<i>Methanocaldococcus jannaschii</i> DSM 2661	yes	yes	yes	yes	yes
<i>Methanocaldococcus vulcanius</i> M7	yes	yes	yes	yes	yes
<i>Methanotorris igneus</i> Kol 5	yes	yes	yes	yes	yes
<b>Methanococci -&gt; Methanococcales -&gt; Methanococcaceae</b>					
<i>Methanococcus aeolicus</i> Nankai 3	yes	yes	yes	yes	yes
<i>Methanococcus maripaludis</i> C5	yes	yes	yes	yes	yes
<i>Methanococcus maripaludis</i> C6	yes	yes	yes	yes	yes
<i>Methanococcus maripaludis</i> C7	yes	yes	yes	yes	yes
<i>Methanococcus maripaludis</i> S2	yes	yes	yes	yes	yes
<i>Methanococcus maripaludis</i> X1	yes	yes	yes	yes	yes
<i>Methanococcus vannielii</i> SB	yes	yes	yes	yes	yes
<i>Methanothermococcus okinawensis</i> IH1	yes	yes	yes	yes	yes
<b>Methanopyri -&gt; Methanopyrales -&gt; Methanopyraceae</b>					
<i>Methanopyrus kandleri</i> AV19	yes	yes	yes	yes	yes
<b>Methanomicrobia -&gt; Methanosarcinales -&gt; Methanosaetaceae</b>					
<i>Methanosaeta concilii</i> GP6	yes	yes	yes	yes	yes
<i>Methanosaeta harundinacea</i> 6Ac	yes	yes	yes	yes	yes
<i>Methanosaeta thermophila</i> PT	yes	yes	yes	yes	yes
<b>Methanomicrobia -&gt; Methanosarcinales -&gt; Methanosarcinaceae</b>					
<i>Methanococcoides burtonii</i> DSM 6242	yes	yes	yes	yes	yes
<i>Methanohalobium evestigatum</i> Z 7303	yes	yes	yes	yes	yes
<i>Methanohalophilus mahii</i> DSM 5219	yes	yes	yes	yes	yes
<i>Methanlobus psychrophilus</i> R15	yes	yes	yes	yes	yes
<i>Methanomethylovorans hollandica</i> DSM 15978	yes	yes	yes	yes	yes
<i>Methanosalsum zhilinae</i> DSM 4017	yes	yes	yes	yes	yes
<i>Methanosarcina acetivorans</i> C2A	yes	yes	yes	yes	yes
<i>Methanosarcina barkeri</i> Fusaro	yes	yes	yes	yes	yes
<i>Methanosarcina mazei</i> Go1	yes	yes	yes	yes	yes
<i>Methanosarcina mazei</i> Tuc01	yes	yes	yes	yes	yes
<b>Crenarchaeota -&gt; Thermoprotei -&gt; Desulfurococcales -&gt; Desulfurococcaceae</b>					
<i>Aeropyrum camini</i> SY1 JCM 12091	yes	yes	yes	yes	yes
<i>Aeropyrum pernix</i> K1	yes	yes	yes	yes	yes
<i>Ignicoccus hospitalis</i> KIN4 I	yes	yes	yes	yes §	yes §
<i>Staphylothermus hellenicus</i> DSM 12710	yes	yes	yes	yes	yes
<i>Staphylothermus marinus</i> F1	yes	yes	yes	yes	yes
<b>Crenarchaeota -&gt; Thermoprotei -&gt; Desulfurococcales -&gt; Pyrodictiaceae</b>					
<i>Hyperthermus butylicus</i> DSM 5456	yes	yes	yes	yes	

<i>Pyrolobus fumarii</i> 1A	yes	yes	yes	yes
<b>Crenarchaeota -&gt; Thermoprotei -&gt; Fervidicoccales -&gt; Fervidicoccaceae</b>				
<i>Fervidicoccus fontis</i> Kam940	yes	yes	yes	(yes)

**Table S1.** Taxonomic overview of putative HARP homologs and RNP RNase P components in other species provided at the NCBI database (downloaded on Oct. 6th, 2015; for RnpA entries, the NCBI database of Oct. 2016 was searched in addition). Taxonomic groups are indicated in bold letters. HARP homologs were searched in two iterations based on the Aq\_880 reference sequence (first round) and all identified sequences presented herein (second round) using blastp+ (at proteome level) and tblastn+ (at genome level) (7) with an *e* value of  $10^{-10}$ . Purely genomic hits (no protein annotation) are indicated by parentheses. To search for RnpA homologs (protein subunit of bacterial RNA-based RNase P, also termed C5) and Rpp29 (also termed Pop4, subunit of archaeal RNA-based RNase P), we used the reference sets provided by (36). The RNA subunit was located using Infernal (30) with the Rfam v12 models of "Bacterial RNase P class A", "B" and "Archaeal RNase P" (31). All results were aligned and manually inspected for false positives. §, retrieved upon increasing the *e* value from  $10^{-10}$  to  $10^{-6}$ ; \*, encoded on a plasmid, not on the main chromosome. HARP: Homolog of *Aquifex* RNase P; RNA: bacterial RNase P RNA subunit; RnpA: bacterial RNase P protein subunit; Rpp29, Pop5, Rpp30: archaeal RNase P protein subunits. We have omitted Rpp21 in our searches because we assumed that Rpp21 is present in Archaea when Rpp29 is present, considering that the two function as a binary complex that binds to the S-domain of archaeal RNase P RNA (43) (49). The same pertains to Pop5/Rpp30, forming a binary complex binding to the C-domain of archaeal RNase P RNA, but here we listed homologs of both because we found Pop5/Rpp30 homologs in HARP-encoding bacteria, prompting us to take a deeper look into the presence of Pop5 and Rpp30. The  $\gamma$ -proteobacterium *Methylohalobius crimeensis* was included based on the recent classification study on the PIN domain-like superfamily by (28).

## Supplementary Materials and Methods

### Purification and enrichment of RNase P activity

#### *Cell lysis using a French press*

A frozen pellet (2 g) of *A. aeolicus* cells harvested at late exponential growth phase was resuspended in 20 ml ice-cold buffer A60 (50 mM Tris-HCl, 60 mM  $\text{NH}_4\text{Cl}$ , 10 mM  $\text{Mg}(\text{OAc})_2$ , 6 mM DTT, pH 7.45) by vigorous vortexing. The homogenous cell suspension was filled into a syringe and pressed three times through a cannula ( $\varnothing = 0.9$  mm). Cells were disrupted by three passages through a French Pressure Cell Press (American Instruments Co.; precooled to  $4^\circ\text{C}$ ) at 1.000 psi. Cell debris was removed by ultracentrifugation at  $125,750 \times g$  in a fixed angle rotor at  $4^\circ\text{C}$  for 1 h (35,000 rpm using a Ti 70 fixed angle rotor). The lysate was stored at  $-80^\circ\text{C}$  until use.

Purification by FPLC was performed with an ÄKTA<sup>TM</sup> basic 10 (GE Healthcare) at room temperature.

#### *Anion exchange chromatography*

Anion exchange chromatography (AEC) was performed using a column packed with Diethylaminoethyl (DEAE) Sepharose Fast Flow (GE Healthcare) to a column volume (CV) of 20 ml. The column was equilibrated with 3 x CV of buffer A60 before 10 ml of filtrated (Filtropur S 0.2 from Sarstedt, Germany) cell lysate was applied to the DEAE column. After washing with 2.25 x CV of buffer A60 the elution was started using a step gradient: at each step, the amount of DEAE elution buffer A1000 (50 mM Tris-HCl, 1 M  $\text{NH}_4\text{Cl}$ , 10 mM  $\text{Mg}(\text{OAc})_2$ , 6 mM DTT, pH 7.45) was increased by 10% at the expense of buffer A60 (100% at the beginning). The flow rate was set to 2 ml/min and the elution volume at each step was 2.2 x CV. At the beginning, 4 ml fractions were collected until step 3. Henceforward, 2 ml fractions were collected. Of each elution fraction, 5  $\mu\text{l}$  were tested for RNase P activity (see below). Fractions with RNase P activity reproducibly eluted in step 3 (30% buffer A1000/70% buffer A60). The fractions with maximum RNase P activity from two independent DEAE purification runs were pooled and concentrated (Amicon® Ultra-15 Ultracel®-10K Centrifugal Filters, Merck Millipore, Germany) at 4°C in a swinging bucket rotor to a final volume of 1.5 - 3 ml. Finally, a 4 M  $(\text{NH}_4)_2\text{SO}_4$  stock solution (pH 7.5) was added to adjust the RNase P sample to a final concentration of 1 M  $(\text{NH}_4)_2\text{SO}_4$ .

#### *Hydrophobic interaction chromatography*

Hydrophobic interaction chromatography (HIC) was performed using a 5 ml HiScreen Phenyl FF low sub column (GE Healthcare). To reduce the loss of sample during passage through a sterile filter (Filtropur S 0.2 from Sarstedt, Germany), HIC start buffer A2 (50 mM Tris-HCl, 60 mM  $\text{NH}_4\text{Cl}$ , 10 mM  $\text{Mg}(\text{OAc})_2$ , 6 mM DTT, 1 M  $(\text{NH}_4)_2\text{SO}_4$ , pH 7.45) was added to the sample to reach a total volume of 10 ml. Before sample loading, the column was equilibrated thoroughly with buffer A2 at a flow rate of 1 - 2 ml/min. After loading, non-interacting sample components were washed out with 3 x CV buffer A2 at a flow rate of 1 - 2 ml/min. Elution was performed with a step gradient using buffer A2 (100% at the beginning) and buffer A60. With each elution step, the proportion of buffer A60 was increased by 20%. The total elution volume of each step was 5 x CV. RNase P activity fractions eluted in step 3 at 0.4 M  $(\text{NH}_4)_2\text{SO}_4$  (40% buffer A2 and 60% buffer A60), in step 4 at 0.2 M  $(\text{NH}_4)_2\text{SO}_4$  (20% HIC buffer A2 and 80% buffer A60) and in step 5 at 0 M  $(\text{NH}_4)_2\text{SO}_4$  (100% buffer A60). The eluted fractions with RNase P activity from each step were pooled separately and designated "HIC<sub>0.2</sub>", "HIC<sub>0.4</sub>" and "HIC<sub>0</sub>". To decrease the salt concentration, fractions HIC<sub>0.2</sub> and HIC<sub>0.4</sub> were dialyzed (ZelluTrans dialysis membranes, MWCO 12 - 14 kDa from Roth, Germany) overnight at 4°C against the 10 to 15-fold sample volume of buffer A60 (three replacements with fresh buffer). For further purification, samples HIC<sub>0.4</sub>, HIC<sub>0.2</sub> and HIC<sub>0</sub> were

concentrated (Amicon® Ultra-15 Centrifugal Filters Ultracel®-10K; Merck Millipore) to final volumes of 500 to 750 µl, respectively. Concentrated HIC<sub>0.4</sub> and HIC<sub>0.2</sub> fractions were further purified separately by size exclusion chromatography (SEC). Ten µl of the concentrated HIC<sub>0</sub> sample were directly analyzed by mass spectrometry (pooled and concentrated fractions from 16 HIC purifications, total volume of 1 ml). 3 µl of the concentrated HIC<sub>0</sub> fraction and 1 - 2 µl of the dialyzed and concentrated HIC<sub>0.4</sub> and HIC<sub>0.2</sub> fractions were tested for RNase P activity in parallel.

#### *Size exclusion chromatography*

HIC<sub>0.4</sub> and HIC<sub>0.2</sub> fractions from two and four separate HIC purifications, respectively, were pooled and concentrated by centrifugal filters (Amicon Ultra-2 mL, Ultracel 10K; Merck Millipore) to final volumes of 125 - 150 µl. Then, the concentrated samples were filtrated through centrifugal filters (Ultrafree-MC, GV 0.22 µm; Merck Millipore) before separate loading on Superose 6 10/300 GL columns (24 ml, GE Healthcare) equilibrated with elution buffer. Elution was performed at a flow rate of 0.3 ml/min with a 1.25-fold column volume (30 ml) of buffer A60 (HIC<sub>0.2</sub> samples) or buffer A260 (HIC<sub>0.4</sub> samples; 50 mM Tris-HCl, 260 mM NH<sub>4</sub>Cl, 10 mM Mg(OAc)<sub>2</sub>, 6 mM DTT, pH 7.45). Eluates were collected in fractions of 0.5 ml. Fractions with RNase P activity eluting from Superose 6 columns loaded with HIC<sub>0.4</sub> or HIC<sub>0.2</sub> samples were designated "SEC<sub>0.4</sub>" and "SEC<sub>0.2</sub>" fractions, respectively. After sample injection, several aliquots were directly taken from the 0.5 ml fractions eluting between 9 and 20 ml. Thereof, 10-16 µl were withdrawn and tested for RNase P activity after supplementation with Mg(OAc)<sub>2</sub> to a final concentration of 10 mM, 40 µl were taken for determining the protein concentration (Bradford assay), 15 µl for SDS-PAGE and 25 µl for mass spectrometry.

#### *Step gradient SDS-PAGE*

Aliquots of SEC fractions containing RNase P activity were analyzed by 20% SDS-PAGE in the order of their elution from the SEC column. According to the strength of RNase P activity, SEC fractions were classified as having low or no activity (+/-), increasing or decreasing RNase P activity (+) or maximum activity (++). As mentioned above, 15 µl and 25 µl aliquots of fractions from each category were analyzed by step gradient SDS-PAGE and mass spectrometry, respectively.

For SDS-PAGE, the discontinuous gel system of Laemmli (24) was used. 4x separation gel buffer (1.5 M Tris-HCl pH 8.8, 0.6 % SDS) was mixed with different amounts of Rotiphorese® Gel 30 (solution of acrylamide/bisacrylamide with 0.8 % bisacrylamide, ratio 37.5 : 1; from Roth, Germany) to obtain gel solutions containing 1x separation gel buffer and

15, 13, 11, 9, 8 or 7% polyacrylamide (PAA). Six separation gel layers of equal height were poured onto each other (15% at the bottom, 7% at the top) after the previous layer had polymerized. Finally, a 4% stacking gel was poured onto the top containing 1x stacking gel buffer (0.5 M Tris-HCl pH 6.8, 0.6 % SDS) and 4% PAA. 15 µl sample aliquots were mixed with 5 µl of 4x protein sample buffer (4% (w/v) SDS, 20% (w/v) glycerol, 5% (w/v) 2-mercaptoethanol, 160 mM Tris-HCl pH 6.8 and 0.04% (w/v) bromophenol blue) and heated to 95°C for 4 min before gel loading. Electrophoresis was performed in 1x running buffer diluted from a 10x running buffer stock (Rotiphorese® 10x SDS-PAGE; Roth, Germany) immediately before electrophoresis. The gel was run until bromophenol blue had almost reached the bottom of the gel. Gels were stained with Coomassie Brilliant Blue G250 as described by (10). Bands of interest were excised with a sterile scalpel and analyzed by mass spectrometry.

#### **Qualitative and quantitative mass spectrometry analysis**

*Probes prepared from gel slices:* After staining the SDS gel with colloidal Coomassie (10), bands of interest were cut out of the gel using a sterile scalpel. After destaining, samples were digested "in-gel" by the addition of Sequencing Grade Modified Trypsin (Promega) overnight at 37°C.

*Liquid probes:* Mass spectrometric analysis of liquid samples was performed with 10 µl of concentrated fraction HIC<sub>0</sub> and 25 µl each of the elution fractions SEC<sub>0.2</sub> and SEC<sub>0.4</sub>. For quantitative (SEC<sub>0.2</sub> and SEC<sub>0.4</sub>) or qualitative (HIC<sub>0</sub>) mass spectrometric analysis the liquid samples were digested with Sequencing Grade Modified Trypsin (Promega) at 37°C overnight.

The mass spectrometric analysis of the samples was performed using an Orbitrap Velos Pro mass spectrometer (Thermo Scientific). An Ultimate nanoRSLC-HPLC system (Dionex), equipped with a custom 20 cm x 75 µm C18 RP column filled with 1.7 µm beads was connected online to the mass spectrometer through a Proxeon nanospray source. Depending on sample concentration, 1 - 15 µl of the tryptic digest were injected onto a C18 pre-concentration column. Automated trapping and desalting of the sample was performed at a flow rate of 6 µl/min using 0.05% formic acid in demineralized and deionized water as solvent.

For separation of the tryptic peptides the following gradient of solvent A (0.05% formic acid in demineralized and deionized water) and solvent B (80% acetonitrile/0.045% formic acid) was performed at a flow rate of 300 nL/min: After holding 4% B for 5 min, a linear gradient to 45% B within 30 min was applied, followed by a linear increase to 95% solvent B within the next 5 min. The column was connected to a stainless steel nanoemitter (Proxeon, Denmark) and the eluent was sprayed directly towards the heated capillary of the mass



spectrometer using a potential of 2300 V. A survey scan with a resolution of 60000 within the Orbitrap mass analyzer was combined with at least three data-dependent MS/MS scans with dynamic exclusion for 30 s, either using CID with the linear ion-trap or HCD combined with orbitrap detection at a resolution of 7500.

Data analysis was performed using Proteome Discoverer (Thermo Scientific) with SEQUEST (11) and Mascot (Matrix Science, version 2.2) search engines, employing either the Swiss-Prot (4) or NCBI (32) databases.

For relative quantification of single proteins in liquid SEC fractions, 50 ng cytochrome *c* was added as internal standard.

#### ***In vitro* transcription of *T. thermophilus* ptRNA<sup>Gly</sup>, *E. coli* RNase P RNA and *B. subtilis* 6S-1 RNA**

All RNAs were synthesized by *in vitro* run-off transcription using T7 RNA polymerase (14) (protocol 1) and linearized plasmid DNAs as template (see Table S2). Transcription products were first subjected to phenol extraction using Roti®-Aqua-Phenol (Roth, Germany), followed by chloroform extraction and isopropanol precipitation for 1-2 h at -20°C (addition of 1 volume isopropanol,  $1/10$  volume 3 M NaOAc (pH 5.0) and  $1/100$  volume 20 mg/ml (w/v) glycogen). After centrifugation (60 min at 13,800 x *g* at 4°C) the supernatant was discarded and the pellet was solved in ddH<sub>2</sub>O after air-drying. Next, the *in vitro* transcript was purified by 8% denaturing (8 M urea) PAGE gels (in 1x TBE buffer) and the RNA band of interest was excised from the gel after UV shadowing using a sterile scalpel. The gel slice was eluted overnight at 8°C under shaking at 900 rpm (Thermomixer, Eppendorf, Germany) in elution buffer (1 M NaOAc, pH 5.0) before the RNA was finally precipitated by isopropanol as described above.

**Table S2:** Plasmids used as template for *in vitro* transcription.

<i>in vitro</i> transcribed RNA	plasmid template	linearization	reference
<i>T. thermophilus</i> ptRNA <sup>Gly</sup>	pSBpt3'hh	Bam HI	(6)
<i>E. coli</i> RNase P RNA	pDW98	Bsa AI	(19)
<i>B. subtilis</i> 6S-1 RNA	pBB-6S-1	Hind III	(5)

**Dephosphorylation, 5'-phosphorylation and 5'-[<sup>32</sup>P]-endlabeling of RNA**

*In vitro* transcripts containing a 5'-triphosphate group were dephosphorylated prior to phosphorylation or 5'-endlabeling. The gel-purified transcripts (1 µg/µl) were incubated in the supplied 1x reaction buffer with 0.3 U/µl thermosensitive alkaline phosphatase (FastAP, Thermo Scientific) at 37°C for 30 min. The enzyme was subsequently inactivated by heating at 75°C for 5 min and the dephosphorylated RNAs were precipitated with ethanol.

RNAs carrying a 5'-monophosphate group were generated by incubation of dephosphorylated RNA (end concentration: 3.5 pmol/µl) in 1x reaction buffer C (50 mM Tris-HCl pH 7.6, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.1 mM spermidine) with 0.5 U/µl T4 polynucleotide kinase (T4 PNK, Thermo Scientific) and 5 mM ATP for 60 min at 37°C. For downstream applications, the 5'-monophosphorylated RNA was extracted first with phenol (Roti®-Aqua-Phenol, Roth), and subsequently with chloroform, precipitated with ethanol and finally dissolved in ddH<sub>2</sub>O. For 5'-[<sup>32</sup>P]-endlabeling, 10-20 pmol dephosphorylated pre-tRNA was incubated with 10 U T4 PNK (Thermo Scientific) and 3 µl [<sup>32</sup>P]-ATP (10 mCi/ml; 3000 Ci/mmol) in 15 µl 1x reaction buffer C (see above) for 60 min at 37°C. The reaction was stopped by addition of 2x RNA sample buffer (2x TBE buffer, 2.6 M urea, 66% deionized formamide, 0.02% [w/v] bromophenol blue and 0.02% [w/v] xylene cyanol blue) and the sample was loaded onto an 8% PAA gel containing 8 M urea for gel purification. After electrophoresis, the radioactively labeled RNA was localized by autoradiography and excised for elution in 1 M NaOAc (pH 5.0) at 8°C overnight under shaking (900 rpm, Thermomixer). After ethanol precipitation the RNA was dissolved in ddH<sub>2</sub>O and stored at -20°C.

**Processing assays***Partially purified A. aeolicus fractions*

The RNase P activity test was performed in a total volume of 20 µl buffer A4.5 (50 mM Tris-HCl pH 7.45, 60 mM NH<sub>4</sub>Cl, 6 mM DTT, 4.5 mM Mg(OAc)<sub>2</sub>) containing 1-16 µl of partially purified cell fraction and trace amounts (< 1 nM, 3,000 Cherenkov cpm ≈ 10,000 dpm) of 5'-[<sup>32</sup>P]-endlabeled ptRNA<sup>Gly</sup> from *Thermus thermophilus*. The cell fractions were preincubated in buffer A4.5 for 5 min at 37°C. The 5'-[<sup>32</sup>P]-endlabeled pre-tRNA substrate was preincubated in parallel in buffer A4.5 for 5 min at 55°C and 25 min at 37°C unless stated otherwise. Immediately after preincubation, the activity assays were started by mixing substrate solution (4 µl) and cell fraction (16 µl) followed by incubation at 37°C for 15 to 60 min. The reactions were stopped by addition of an equal volume 2x RNA sample buffer (see above) and reaction products were separated by 20% denaturing (8 M urea) PAGE. After electrophoresis, substrate and cleavage products were visualized and quantified by phosphorimaging using a Bio-Imaging Analyzer FLA3000-2R (Fujifilm) and the analysis software AIDA (Raytest) as described (50).



*Activity assay with recombinant Aq\_880 variants*

Enzymatic activity of *A. aeolicus* Aq\_880 and mutants thereof (see Fig. S19 and S20) was analyzed in buffer F (50 mM Tris-HCl pH 7.0, 20 mM NaCl, 4.5 mM MgCl<sub>2</sub>, 20 µg/ml BSA, 5 mM DTT; DTT was always freshly added immediately before buffer use) at 37°C under single-turnover conditions. Enzyme and substrate solutions were preincubated separately, enzyme for 5 min at 37°C and substrate for 5 min at 55°C and 25 min at 37°C, both in buffer F. The final concentrations were 50 nM enzyme and < 1 nM substrate (5'-[<sup>32</sup>P]-end-labeled ptRNA<sup>Gly</sup>).

*Determination of single-turnover kinetic parameters for pre-tRNA<sup>Gly</sup> processing by Aq\_880*

Trace amounts (< 1 nM) of 5'-[<sup>32</sup>P]-end-labeled *Thermus thermophilus* pre-tRNA<sup>Gly</sup> were preincubated in buffer F (50 mM Tris-HCl pH 7.0, 20 mM NaCl, 4.5 mM MgCl<sub>2</sub>, 20 µg/ml BSA, 5 mM DTT; DTT was always freshly added immediately before buffer use) for 5 min at 55°C and 25 min at 37°C. Aq\_880 (varied between final concentrations of 5 and 100 nM) was preincubated separately for 5 min at 37°C in buffer F. Reactions were started by combining 16 µl enzyme and 4 µl substrate mix, and aliquots of 4 µl were withdrawn at different time points, such that the first three points were below 70% substrate conversion; reactions were stopped by adding an equal volume 2x RNA sample buffer. Samples were subjected to 20% denaturing (8 M urea) PAGE. The 5'-[<sup>32</sup>P]-labeled pre-tRNA substrate and the 5'-cleavage product were quantified with a Bio-Imaging Analyzer FLA3000-2R (Fujifilm) and the analysis software AIDA (Raytest). First-order rate constants of cleavage ( $k_{\text{obs}}$ ) were calculated by nonlinear regression analysis (Grafit 5.0.13) fitting the data to the equation  $f_{\text{cleaved}} = f_{\text{endpoint}} (1 - e^{-(k_{\text{obs}})t})$ , where  $f_{\text{cleaved}}$  = fraction of pre-tRNA cleaved,  $t$  = time,  $f_{\text{endpoint}}$  = maximum cleavable fraction of pre-tRNA; based on mean  $k_{\text{obs}}$  values derived from 3-6 independent experiments for each Aq\_880 concentration, the kinetic parameters  $K_{\text{m(sto)}}$  and  $k_{\text{react}}$  were calculated by fitting the data to the equation  $k_{\text{obs}} = k_{\text{react}} \times [\text{Aq\_880}] / (K_{\text{m(sto)}} + [\text{Aq\_880}])$ ;  $K_{\text{m(sto)}}$  is the enzyme concentration at the half-maximal rate of cleavage and  $k_{\text{react}}$  is the maximum single-turnover rate (6).

*Determination of multiple-turnover kinetic parameters for pre-tRNA<sup>Gly</sup> processing by Aq\_880*

Experiments were performed as for the single turnover, except that the final Aq\_880 concentration was kept constant at 1 nM and the substrate was varied between 5 and 200 nM.

*E. coli* and *B. subtilis* RNase P holoenzyme and *E. coli* P RNA-alone assays

Enzymatic activity of *E. coli* and *B. subtilis* RNase P holoenzymes (final concentrations: 4 - 50 nM) was analyzed in buffer F (see above) at 37°C unless stated otherwise. For RNase P holoenzyme assays, the P RNA was preincubated for 5 min at 55°C and 5-30 min at 37°C in buffer F followed by incubation with P protein for 5 min at 37°C to reconstitute the holoenzyme. 5'-[<sup>32</sup>P]-end-labeled substrate was preincubated in parallel in the same buffer (5 min at 55°C and 25 min at 37°C) before mixing 16 µl of enzyme and 4 µl of substrate solution to start the reaction.

For *E. coli* P RNA-alone reactions in Fig. S2, S4, S7 and S9, 4.5 pmol *E. coli* P RNA were incubated in 16 µl buffer A0 (50 mM Tris-HCl pH 7.45, 60 mM NH<sub>4</sub>Cl, 6 mM DTT) supplemented with 10 mM Mg(OAc)<sub>2</sub> for 5 min at 37°C. In parallel, trace amounts of 5'-[<sup>32</sup>P]-end-labeled pre-tRNA<sup>Gly</sup> (4 µl) were preincubated for 5 min at 55°C and subsequently for 25 min at 37°C under the same conditions. Reactions were started by combining 16 µl enzyme and 4 µl substrate mix. After 30-60 min of incubation at 37°C, the reaction was stopped by adding an equal volume of 2x RNA sample buffer.

*Thermostability test for Aq\_880 and E. coli* RNase P

Aq\_880 (50 nM) or *E. coli* RNase P holoenzyme (50 nM) were preincubated for 10 min at 85°C in buffer F (50 mM Tris-HCl pH 7.0, 20 mM NaCl, 4.5 mM MgCl<sub>2</sub>, 20 µg/ml BSA, 5 mM DTT; DTT was always freshly added immediately before buffer use), followed by slow (1.8°C/min) cooling to 37°C and incubation for another 5 min at 37°C. In parallel, trace amounts (< 1 nM) of 5'-[<sup>32</sup>P]-end-labeled pre-tRNA<sup>Gly</sup> from *T. thermophilus* were preincubated for 5 min at 55°C and 25 min at 37°C under the same buffer conditions. Processing reactions were started by combining 16 µl enzyme solution and 4 µl substrate solution. After incubation for 30-45 min at 37°C, the reaction was stopped by addition of an equal volume 2x RNA sample buffer. Reaction products were separated by 20% denaturing (8 M urea) PAGE and substrate and cleavage product were visualized and quantified by phosphorimaging as described above.

*Pretreatment with micrococcal nuclease (MN)*

The processing activity of different recombinant RNase P enzymes was investigated after treatment with micrococcal nuclease (MN). The analyzed enzymes and their final concentrations in the activity assay were 550 nM *A. thaliana* PRORP1, 125 nM *A. aeolicus* Aq\_880, 50 nM *E. coli* holoenzyme and 150 nM *E. coli* P RNA. Additionally, the effect of MN treatment on a mixture of 150 nM *E. coli* P RNA plus 125 nM *A. aeolicus* Aq\_880 was tested. The RNase P enzyme/enzyme mix was incubated in 17 µl buffer G (50 mM Tris-HCl pH 7.0, 20 mM NaCl, 4.12 mM CaCl<sub>2</sub>, 20 µg/ml BSA, 5 mM DTT) supplemented with MgCl<sub>2</sub> and

either 1  $\mu$ l MN (Thermo Scientific, 300 U/ $\mu$ l) or 1  $\mu$ l ddH<sub>2</sub>O (control) for 20 min at 37°C/800 rpm in a thermomixer (Eppendorf, Germany). The final Mg<sup>2+</sup> concentration in the cleavage reaction was 4.5 mM for *A. thaliana* PRORP1 and *A. aeolicus* Aq\_880, 10 mM for the *E. coli* holoenzyme and 100 mM for *E. coli* P RNA. The MN activity was stopped by adding 3  $\mu$ l 0.2 M EGTA (pH 8.0) to the enzyme mix. Prior to reaction start, the enzyme mix was incubated for another 5 min at 37°C. In parallel to MN treatment of the enzyme solution, 7 nM of 5'-[<sup>32</sup>P]-end-labeled ptRNA<sup>Gly</sup> from *T. thermophilus* supplemented with 3.5  $\mu$ M carrier RNA (*in vitro* transcribed *B. subtilis* 6S-1 RNA) were preincubated for 5 min at 55°C and 20 min at 37°C in buffer F (see above). The reaction was started by combining 16  $\mu$ l enzyme and 4  $\mu$ l substrate solution. After incubation for 45 min at 37°C the reaction was stopped by addition of an equal volume 2x RNA sample buffer and the enzymatic activity was determined as described above.

#### Cell lysate supplemented with recombinant proteins

The influence of recombinant *A. aeolicus* proteins Aq\_707, Aq\_221 (PNPase) and Aq\_880 on pre-tRNA<sup>Gly</sup> cleavage by the endogenous RNase P activity of partially purified *A. aeolicus* cell lysates (Fig. 2A and S13) was analyzed as follows: recombinant Aq\_707 (final concentration: 2.5  $\mu$ M), Aq\_221 (final concentration: 1.5  $\mu$ M) or Aq\_880 (final concentration: 4.2  $\mu$ M) were combined with HIC<sub>0.4</sub> eluate (3.7  $\mu$ g protein; see Fig. S3) in a volume of 20  $\mu$ l and preincubated for 5 min at 37°C in buffer A0 (50 mM Tris-HCl pH 7.45, 60 mM NH<sub>4</sub>Cl, 6 mM DTT) supplemented with 4.5 mM Mg(OAc)<sub>2</sub>. In parallel, substrate mix containing 164 nM 5'-phosphorylated and trace amounts of 5'-[<sup>32</sup>P]-end-labeled pre-tRNA<sup>Gly</sup> in a total volume of 7  $\mu$ l buffer A0 supplemented with 4.5 mM Mg(OAc)<sub>2</sub> was preincubated for 5 min at 55°C followed by 25 min at 37°C. The reaction was started by combining 16  $\mu$ l preincubated enzyme and 4  $\mu$ l preincubated substrate solution (final substrate concentration: 33 nM). Reaction aliquots of 4  $\mu$ l were taken at different time points (1, 5, 10, 30 and 60 min) and were immediately added to an equal volume of 2x RNA sample buffer to stop the cleavage reaction. The enzymatic activity was analyzed by 20% denaturing (8 M urea) PAGE as described above. As control activity, preassembled (45) *B. subtilis* RNase P holoenzyme (10 nM) was analyzed in KN buffer supplemented with 4.5 mM Mg<sup>2+</sup>.

#### Aq\_880 supplemented with total RNA

The activity of recombinant Aq\_880 (final concentration: 50 nM) was analyzed in the presence of total RNA from *A. aeolicus* (final concentration 20 ng/ $\mu$ l; prepared by acetic acid plus phenol/chloroform extraction or phenol/chloroform extraction; see below) at 37°C or 70°C. For performing the processing assay at 37°C, Aq\_880 was preincubated for 5 min at 37°C in the presence of total RNA in buffer F. For performing the processing assay at 70°C,

the total RNA was preincubated 5 min at 70°C in buffer F before Aq\_880 was added; thereafter, the mixture was incubated for another 10 min at 70°C. In parallel, < 1 nM 5'-[<sup>32</sup>P]-end-labeled pre-tRNA<sup>Gly</sup> was incubated for 10 min at 70°C under the same buffer conditions. The reactions were started by combining preincubated enzyme and substrate solution. After 1, 2, 5 and 30 min incubation at 37°C or 70°C, aliquots were taken and analyzed by 20% denaturing PAGE as described before.

#### Activity assay with other bacterial RNase P RNAs

Enzymatic activity of chimeric holoenzymes composed of bacterial P RNA from *Thermodesulfatator indicus* or *E. coli* (= positive control), and P protein from *B. subtilis* was analyzed in buffer KN (20 mM HEPES-KOH, pH 7.4; 150 mM NH<sub>4</sub>OAc; 2 mM spermidine; 50 μM spermine; 4 mM β-mercaptoethanol) supplemented with 4.5 mM MgCl<sub>2</sub> at 25°C. First, 0.2 pmol bacterial P RNA was preincubated in a total volume of 15 μl for 5 min at 55°C and 15 min at 25°C before addition of 1 μl (1 pmol) *B. subtilis* P protein. For holoenzyme assembly, the sample was incubated for another 5 min at 25°C. For RNA-alone reactions, 0.5 pmol bacterial P RNA was preincubated in a total volume of 16 μl for 5 min at 55°C and 20 min at 25°C in buffer KN supplemented with 100 mM MgCl<sub>2</sub>. In parallel, trace amounts (< 1 nM) of 5'-[<sup>32</sup>P]-end-labeled ptRNA<sup>Gly</sup> from *T. thermophilus* were preincubated in a total volume of 10 μl for 5 min at 55°C and 20 min at 25°C under the same buffer conditions (KN buffer plus 4.5 or 100 mM Mg<sup>2+</sup>). Reaction was started by combining 16 μl preincubated enzyme and 4 μl preincubated substrate mix. Aliquots were taken (after 20 s, 1 min, 5 min and 12 min) and analyzed by 20% denaturing PAGE.

#### Activity assays using other HARP proteins

Activity of HARP proteins was analyzed in buffer F (50 mM Tris-HCl pH 7.0, 20 mM NaCl, 4.5 mM MgCl<sub>2</sub>, 20 μg/ml BSA, 5 mM DTT) using trace amounts (< 1 nM) of 5'-[<sup>32</sup>P]-end-labeled *T. thermophilus* pre-tRNA<sup>Gly</sup>. The latter was preincubated for 5 min at 55°C and 5 min at 37°C. The HARP protein (final concentration 50 nM) was preincubated in parallel in the same buffer for 5 min at 37°C. The reaction was started by combining preincubated enzyme and substrate solutions. Aliquots were taken after 5, 15 and 30 min. Processing reactions were analyzed by 20% denaturing PAGE (see example in Fig. S21).

#### **Construction of bacterial expression vectors for Aq\_880**

For the expression of an Aq\_880 variant containing a C-terminal His tag (= pET28a(+)\_aq880cHis), the *aq\_880* gene was amplified from chromosomal DNA of *A. aeolicus* using the primer pair 630 (5'-AAG CCA TGG ATG TGT TCG TTC TCG ACA C-3'; Nco I site in italics) / 631 (5'-CTC TCG AGA AAC CTG TGT CTT ACC AAG CTC TC-3'; Xho

I site in italics). The chromosomal DNA was prepared with the DNeasy Tissue Kit (Qiagen) according to the instructions of the manufacturer and the PCR amplification was performed with the Long PCR Enzyme Mix (Thermo Scientific). The DNA product was inserted via the primer-encoded restriction sites Nco I and Xho I into vector pET28a(+) (Novagen). The resulting plasmid pET28a(+)\_aq880cHis was transformed into *E. coli* Rosetta (DE3) cells. For the *in vivo* complementation studies in *E. coli* strain BW (48) the expression vector pDG\_aq880cHis was constructed. Therefore, the aq\_880 gene including the sequences for the C-terminal His tag and the T7 terminator sequence at the 3'-end was PCR-amplified from expression vector pET28a(+)\_aq880cHis using the primer pair 645 (5'-AGC CCG GGC ATG GAT GTG TTC GTT CTC GAC AC-3'; Sma I site in italics) / 646 (5'-AGG CAT GCA TCC GGA TAT AGT TCC TCC TTT CAG-3'; Sph I site in italics). The PCR fragment was inserted via the primer-encoded restriction sites Sma I and Sph I into vector pDG148(S/X) (13). In addition, for all Aq\_880 variants, pDG148(S/X) derivatives were constructed that encoded the proteins without His-tag. For this purpose, primer 646 was replaced with primer 054a\_Aq\_pDG\_no His (5'- TTT GCA TGC **TTA** AAA CCT GTG TCT TAC CAA GC-3'; stop codon in bold, Sph I site in italics).

#### Construction of expression vectors for Aq\_880 variants D160A, D138A, D144A and D142A

Plasmids for the production of Aq\_880 variants containing unique aspartate-to-alanine mutations were produced from vector pET28a(+):aq880cHis by site-directed mutagenesis according to the protocol of the QuikChange® XL Site-Directed Mutagenesis Kit (Agilent Technologies). Yet, for PCR reactions we used the *Pwo* DNA polymerase (Pqlab). The primers used for the *in vitro* mutagenesis are listed in Table S3 below. The Aq\_880 variants were expressed and purified as described below for the Aq\_880 wild-type protein under "Expression and purification of recombinant proteins".

	Sequence 5' --> 3'
<b><i>In vitro mutagenesis</i></b>	
009_Aq880_D142A_for	AATACTCGACAGTAAAGAGG <b>CC</b> GTGGATGTGCTGCTCCTTGCC
010_Aq880_D142A_rev	GGCAAGGAGCAGCACATCCACG <b>GC</b> CTCTTTACTGTCTGAGTATT
011_aq880_D160A_for	ACGGGGTTCTCGTTTCGGCGG <b>CT</b> GAAGGCCTCAGAACATGGGC
012_Aq880_D160A_rev	GCCCATGTTCTGAGGCCTTCAG <b>CC</b> CGCCGAAACGAGAACCCCG
013_Aq880D142144A_f	AATACTCGACAGTAAAGAGG <b>CC</b> GTGG <b>CT</b> GTGCTGCTCCTTGCC
014_Aq880_D142144A_r	GGCAAGGAGCAGCACAG <b>CC</b> CACG <b>GC</b> CTCTTTACTGTCTGAGTATT



024_Aq880D138A_for	TCAGGCAGGGAATACTCGCCAGTAAAGAGGACGTGG
025_Aq880D138A_rev	CCACGTCCTCTTTACTGCGAGTATTCCTGCCTGA
026_Aq880D144A_for	AATACTCGACAGTAAAGAGGACGTGGCTGTGCTGCTCCTTGCC
027_Aq880D144A_rev	GGCAAGGAGCAGCACAGCCACGTCCTCTTTACTGTCGAGTATT
<b>complementation</b>	
645 (forward), Sma I site in italics	AGC CCG GGC ATG GAT GTG TTC GTT CTC GAC AC
646 (reverse), for C-His variants, Sph I site in italics	AGG CAT GCA TCC GGA TAT AGT TCC TCC TTT CAG
054a_Aq_pDG_no His (reverse), stop codon in bold, Sph I site in italics	TTT GCA TGC <b>TTA</b> AAA CCT GTG TCT TAC CAA GC

**Table S3:** Oligonucleotides used for site-directed mutagenesis and complementation analyses in *E. coli* BW. Codon mutations are indicated by bold letters, with the altered codons underlined.

#### Expression and purification of recombinant proteins

Aq\_880 variants carried a C-terminal His tag for purification by Ni-NTA affinity chromatography. Genes for the recombinant proteins were inserted into the expression vector pET28a(+) (Novagen) and were expressed in *E. coli* strain Rosetta (DE3). The Aq\_880 variants were either prepared by "batch" purification or using the ÄKTApurifier chromatography system (GE Healthcare). Generally, 250-800 ml of "auto induction" LB medium supplemented with 0.05% (w/v) glucose and 0.2% (w/v) lactose (40) were inoculated with a single colony followed by growth overnight at 37°C/200 rpm in an air shaker (GFL, model 3033, Germany). After cell harvest by centrifugation, the cell pellet was resuspended in 10 ml (or 30 ml for ÄKTA purification) NPI-10 buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM Imidazol) and bacteria were lysed by sonication. Cell debris was pelleted by centrifugation for 45 min at 8.000 x g/4°C. For "batch" purification 1 ml of Protino® Ni-NTA (Machery-Nagel) was equilibrated with NPI-10 buffer before adding the cleared cell lysate. After rotation overhead for 60-120 min at 4°C the slurry was centrifugated for 2 min at 2,500 x g/4°C, the supernatant was discarded and the matrix was washed with 1 ml buffer NPI-10 (4-6x). For elution, the matrix was incubated 3 to 5 times with 1 ml NPI-250 buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 250 mM Imidazol). The eluted fractions were pooled and dialyzed (MWCO 12-14 kDa) against storage buffer (10 mM Tris-HCl pH 8, 100 mM KCl, 0.1 mM EDTA, 3 mM DTT, 50% glycerin, and 10 mM MgCl<sub>2</sub> if required). For ÄKTA purification, a 1 ml HisTrap HP (GE Healthcare) column was used. After equilibration with NPI-10 buffer the cleared lysate was loaded onto the column and washed with NPI-10 buffer until the flow

through reached baseline absorption at 254 and 280 nm. Proteins were finally eluted with NPI-250 buffer collecting 1-2 ml fractions. The purity of the preparation was controlled by SDS-PAGE and Coomassie brilliant blue staining. Elution fractions containing the protein were pooled and dialyzed against storage buffer (see above).

The bacterial RNase P proteins from *E. coli* and *B. subtilis* were purified using a N-terminal polyhistidine affinity tag (His-tagged peptide leader: MRGSHHHHHGS). The His-tagged proteins are encoded in plasmid pQE-30 and were expressed in *E. coli* strain JM109. Expression and purification were performed as described previously (35) (48).

The PRORP1 gene from *A. thaliana* was cloned into the NcoI/XhoI sites of vector pET28b(+) (Novagen) and was expressed in *E. coli* BL21(DE3). The recombinant variant of PRORP1 lacked the mitochondrial targeting sequence of the native precursor and contained a C-terminal His-tag for affinity chromatography purification as described in (33).

For activity assays, Aq\_880 and PRORP1 were diluted in enzyme dilution buffer (30 mM Tris pH 7.8, 30 mM NaCl, 10 % (v/v) glycerol, 10 µg/ml BSA, 1 mM DTT, 0.3 mM EDTA).

#### RT-PCR analysis

First, the detection limit of *E. coli* P RNA in the RT-PCR assay (Access RT-PCR system kit, Promega) was determined. Different amounts of *in vitro* transcribed *E. coli* P RNA (10 amol, 1 amol, 100 zmol and 10 zmol) were used as template for RT-PCR-amplification with specific primers EcM1-5 (5-GAA GCT GAC CAG ACA GTC GCC GCT TC-3) and EcM1-3 (5-ATC TAG GCC AGC AAT CGC TC-3). *In vitro* transcribed *B. subtilis* 6S-1 RNA (20 pmol) was added to each reaction as carrier RNA. The RT-PCR reaction was performed according to the manufacturer's protocol. Next, different fractions from the preparation of recombinant Aq\_880 and Aq\_880[D142] were tested for *E. coli* P RNA contamination, including fractions from "batch" and ÄKTA purification: (i) elution fractions after dialysis (each 1.25 pmol protein; the same amount that was used in the activity assays); (ii) flow through fractions (each 2.5 µl; unbound sample washed out before elution) and (iii) cell lysates (each 2.5 µl; soluble cell fractions after sonication). Each sample was filled with ddH<sub>2</sub>O to a final volume of 400 µl and was subjected to phenol extraction (Roti®-Aqua-Phenol, Roth) followed by chloroform extraction. After subsequent ethanol precipitation, the air-dried RNA pellet was resolved in 10 µl ddH<sub>2</sub>O for RT-PCR analysis (see above). In additional control reactions the reverse transcriptase was omitted. (RT-)PCR products (each 3 µl) were separated on 2% agarose gels and stained with ethidium bromide.

**A. *aeolicus* total RNA preparation***Phenol-chloroform extraction*

Total RNA was prepared according to Method 1 ("Extracting RNA three times with hot phenol.") described by (8) using 0.1 g cell pellet from *A. aeolicus*. The lysozyme treatment was omitted.

*Acetic acid extraction*

Total RNA was prepared from RNase P-enriched *A. aeolicus* HIC<sub>0.2</sub> and HIC<sub>0.4</sub> fractions by the "acetic acid extraction" method originally developed for the separation of ribosomal RNA and proteins (17) (18). Shortly,  $\frac{1}{10}$  volume icecold 1 M Mg(OAc)<sub>2</sub> and 2 volumes icecold acetic acid were added to the concentrated (Amicon®Ultra 10K, Millipore) *A. aeolicus* cell fraction (130-150  $\mu$ l) and incubated for 45 min on ice in a shaking device at 200 rpm. Next, RNA and protein were separated by centrifugation for 30 min at 9,600  $\times g$  / 4°C. The protein-containing supernatant was carefully removed from the pelleted RNA. The RNA pellet was resuspended in the 20-fold volume (relative to the volume of initial RNase P-enriched cell fraction) of buffer AAE/urea (20 mM Tris-HCl pH 7.4, 4 mM Mg(OAc)<sub>2</sub>, 400 mM NH<sub>4</sub>Cl, 0.16 mM EDTA, 6 M urea, 4 mM DTT) and dialyzed against the 10,000-fold volume of AAE/urea buffer at 4°C overnight. Subsequently, the RNA fraction was dialyzed twice against the 5,000-fold volume of urea-free AAE buffer for 45 min. The dialysis was immediately stopped when clouding in the dialysis bag was observed. After dialysis the RNA-containing fraction was first centrifugated for 30 min at 4,000  $\times g$  / 4°C and then concentrated to the 1.5-fold volume of the initial RNase P-enriched cell fraction using centrifugal filters. Finally, the RNA was precipitated with ethanol, washed twice with 70% ethanol and resuspended in buffer A60 (50 mM Tris-HCl pH 7.45, 60 mM NH<sub>4</sub>Cl, 10 mM Mg(OAc)<sub>2</sub>, 6 mM DTT). For the experiment shown in Fig. S14, the total RNA was further subjected to phenol extraction (Roti®-Aqua-Phenol, Roth), chloroform extraction and ethanol precipitation.

**Size exclusion chromatography of recombinant Aq\_880**

First, the Superose 6 10/300 GL column (24 ml, GE Healthcare) was calibrated using Apoferritin [from horse spleen] (Sigma-Aldrich) and the Gel Filtration Markers Kit (Sigma-Aldrich). After equilibration with buffer S (50 mM Tris-HCl pH 7.45, 150 mM NaCl, 4.5 mM Mg(OAc)<sub>2</sub>, 5 mM DTT) the standard proteins (bovine serum albumin, alcohol dehydrogenase,  $\beta$ -amylase, carbonic anhydrase, cytochrome c and apoferritin; each 0.6 mg) were loaded onto and passed through the column at a flow rate of 0.3 ml/min at room temperature. The elution volume of each standard protein was determined. To find out whether Aq\_880 forms oligomers, 1.3 mg recombinant Aq\_880 (ÅKTA-purified) was applied to the calibrated column as described for the standard proteins. Before application, the Aq\_880 was filtrated



(Ultrafree-MC, GV 0.22 µm, Merck Millipore). The oligomerization state of Aq\_880 was inferred from comparing the elution volume of recombinant Aq\_880 with the elution volumes of the standard proteins.

#### Complementation studies

In the *E. coli* strain BW, expression of the essential *mnpB* gene (encoding the RNase P RNA subunit) is under control of an arabinose-inducible promoter (48). Under non-permissive conditions the survival of the cells depends on a functional plasmid-encoded RNase P activity. For *in vivo* analysis, the different pDG148-based expression vectors were introduced into *E. coli* BW cells by transformation of chemically competent cells. For the complementation tests, transformed cells were plated on LB agar plates supplemented with 10 mM arabinose, 100 µg/ml ampicillin and 34 µg/ml chloramphenicol. After incubation overnight at 37°C, single colonies were resuspended in 500 µl LB medium, pelleted in a desktop centrifuge and again resuspended in 500 µl LB medium. 3 to 5 aliquots (each 10 µl) were dropped in parallel onto selective LB agar plates supplemented either with 10 mM arabinose (permissive conditions) or 10 mM glucose (non-permissive conditions). The plates were incubated for 1 to 5 days at 37°C or 30°C. Other HARP genes were analyzed in the same way using gene-specific primers for the construction of corresponding pDG148(S/X) derivatives.

For the complementation studies in *Saccharomyces cerevisiae*, *aq\_880* variants were constructed and amplified by PCR techniques and then cloned into the XbaI/HindIII sites of a derivative of plasmid YEplac181 containing the truncated *S. cerevisiae ADH1* promoter exactly as previously described (47). The previously described plasmid-shuffle system (47) was used to test whether *aq\_880* (or its variants) enabled the tester strain with its chromosomal *RPR1* deletion (*rpr1 Δ::kanMX4*) to lose the *RPR1* expression plasmid with the *URA3* marker and thereby survive selection on 5-fluoroorotic acid. Colonies obtained through rescue by Aq\_880 or its variant D144A were isolated, regrown, and subsequently applied as spots in 10-fold serial dilution to YPD plates. The growth of the strains was monitored in parallel to a control (rescue by *RPR1*). PCR genotyping was carried out as previously described (42). To verify the presence of *aq\_880*, primers flanking the gene in the vector were used (primer *for*, 5'-CCTCGTCATTGTTCTCGTT; primer *rev*, 5'-GGCTCGTATGTTGTGTGGA). All other genotyping primers were described previously (47).

#### A. *aeolicus* proteins identified by mass spectrometry - occurrence of homologs in the *Aquificales*

Analyses for the presence of Aq\_880 and Aq\_707 homologs in the *Aquificales* and particularly in the *Aquificaceae* were performed with BLAST+ (7; *e* value 10<sup>-5</sup>) and the

genomes published in the NCBI database (34; download 2013-08-05). The classification of the organisms was based on the NCBI taxonomy database.

## References

- (1) Anantharaman V, Aravind L (2006) The NYN domains: novel predicted RNases with a PIN domain-like fold. *RNA biology* 3(1):18-27.
- (2) Arcus VL, McKenzie JL, Robson J, Cook GM (2010) The PIN-domain ribonucleases and the prokaryotic VapBC toxin-antitoxin array. *Protein Eng. Des. Sel.* 24(1-2):33-40. doi: 10.1093/protein/gzq081.
- (3) Andrews AJ, Hall TA, Brown JW (2001) Characterization of RNase P holoenzymes from *Methanococcus jannaschii* and *Methanothermobacter thermoautotrophicus*. *Biol. Chem.* 382(8):1171-1177. doi: 10.1515/BC.2001.147.
- (4) Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 28(1):45-48.
- (5) Beckmann BM, Burenina OY, Hoch PG, Kubareva EA, Sharma CM, Hartmann RK (2011) In vivo and in vitro analysis of 6S RNA-templated short transcripts in *Bacillus subtilis*. *RNA Biol.* 8(5):839-849. doi: 10.4161/rna.8.5.16151.
- (6) Busch S, Kirsebom LA, Notbohm H, Hartmann RK (2000) Differential role of the intermolecular base-pairs G292-C(75) and G293-C(74) in the reaction catalyzed by *Escherichia coli* RNase P RNA. *J. Mol. Biol.* 299(4):941-951. doi: 10.1006/jmbi.2000.3789.
- (7) Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2008) BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421.
- (8) Damm K, Bach S, Müller KM, Klug G, Burenina OY, Kubareva EA, Grünweller A, Hartmann RK (2015) Impact of RNA Isolation Protocols on Detection by Northern Blotting. *Methods Mol. Biol.* 1296:29-38. doi: 10.1007/978-1-4939-2547-6\_4.
- (9) Das R, Loss S, Li J, Waugh DS, Tarasov S, Wingfield PT, Byrd RA, Altieri AS (2008) Structural biophysics of the NusB:NusE antitermination complex. *J. Mol. Biol.* 376(3):705-720. doi: 10.1016/j.jmb.2007.11.022.
- (10) Dyballa N, Metzger S (2009) Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *J. Vis. Exp.* pii: 1431. doi: 10.3791/1431.
- (11) Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 5(11):976-989. doi: 10.1016/1044-0305(94)80016-2.

- (12) Gough J, Karplus K, Hughey R, Chothia C (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* 313(4):903-919. doi: 10.1006/jmbi.2001.5080.
- (13) Gößringer M, Hartmann RK (2007) Function of heterologous and truncated RNase P proteins in *Bacillus subtilis*. *Mol. Microbiol.* **66**(3):801-813. doi: 10.1111/j.1365-2958.2007.05962.x.
- (14) Gößringer M, Helmecke D, Köhler K, Schön A, Kirsebom LA, Bindereif A, Hartmann RK (2014) Enzymatic RNA Synthesis Using Bacteriophage T7 RNA Polymerase. *Handbook of RNA Biochemistry Vol. 1*, eds Hartmann RK, Bindereif A, Schön A, Westhof E) (Wiley-VCH), pp 21-27.
- (15) Hall TA, Brown JW (2002) Archaeal RNase P has multiple protein subunits homologous to eukaryotic nuclear RNase P proteins. *RNA* 8(3):296-306.
- (16) Hall TA, Brown JW (2004) Interactions between RNase P protein subunits in archaea. *Archaea* 1(4):247-254.
- (17) Hardy SJ, Kurland CG, Voynow P, Mora G (1969) The ribosomal proteins of *Escherichia coli*. I. Purification of the 30S ribosomal proteins. *Biochemistry*. 8(7):2897-28905.
- (18) Hartmann RK, Vogel DW, Walker RT, Erdmann VA (1988) In vitro incorporation of eubacterial, archaeobacterial and eukaryotic 5S rRNAs into large ribosomal subunits of *Bacillus stearothermophilus*. *Nucleic Acids Res.* 16(8):3511-24.
- (19) Heide C, Pfeiffer T, Nolan JM, Hartmann RK (1999) Guanosine 2-NH<sub>2</sub> groups of *Escherichia coli* RNase P RNA involved in intramolecular tertiary contacts and direct interactions with tRNA. *RNA*. 5(1):102-116.
- (20) Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* 135(3):462-474. doi: 10.1016/j.cell.2008.09.013.
- (21) Howard MJ, Lim WH, Fierke CA, Koutmos M (2012) Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proc. Natl. Acad. Sci. U S A.* 109(40):16149-16154. doi: 10.1073/pnas.1209062109.
- (22) Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10(6):845-858. doi: 10.1038/nprot.2015.053.

- (23) Kitamura A, Nishimoto M, Sengoku T, Shibata R, Jäger G, Björk GR, Grosjean H, Yokoyama S, Bessho Y (2012) Characterization and structure of the *Aquifex aeolicus* protein DUF752: a bacterial tRNA-methyltransferase (MnmC2) functioning without the usually fused oxidase domain (MnmC1). *J. Biol. Chem.* 287(52):43950-43960. doi: 10.1074/jbc.M112.409300.
- (24) Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227(5259):680-685.
- (25) Lechner M, Nickel AI, Wehner S, Riege K, Beckmann BM, Hartmann RK, Marz M (2014) Genomewide comparison and novel ncRNAs of *Aquificales*. *BMC Genomics* 15:522. doi: 10.1186/1471-2164-15-522.
- (26) Marszałkowski M, Teune JH, Steger G, Hartmann RK, Willkomm DK (2006) Thermostable RNase P RNAs lacking P18 identified in the *Aquificales*. *RNA* 12(11):1915-1921. doi: 10.1261/rna.242806.
- (27) Marszałkowski M, Willkomm DK, Hartmann RK (2008) 5'-end maturation of tRNA in *Aquifex aeolicus*. *Biol. Chem.* 389(4):395-403. doi: 10.1515/BC.2008.042.
- (28) Matelska D, Steczkiewicz K, Ginalski K (2017) Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Res.* 2017 May 31. doi: 10.1093/nar/gkx494. [Epub ahead of print].
- (29) Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247(4):536-540. doi: 10.1006/jmbi.1995.0159
- (30) Nawrocki EP, Eddy SR (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 29(22):2933-2935. doi: 10.1093/bioinformatics/btt509.
- (31) Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, Finn RD (2015) Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* 43 (Database issue), D130–D137 (2015). doi: 10.1093/nar/gku1063.
- (32) NCBI Resource Coordinators (2014) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 42, D7-17. doi: 10.1093/nar/gkt1146.
- (33) Pavlova LV, Gößringer M, Weber C, Buzet A, Rossmann W, Hartmann RK (2012) tRNA processing by protein-only versus RNA-based RNase P: kinetic analysis reveals mechanistic differences. *ChemBiochem* 13(15):2270-2276. doi: 10.1002/cbic.201200434.

- (34) Pruitt KD, Tatusova T, Brown GR, Maglott DR (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* 40 (Database issue):D130-5. doi: 10.1093/nar/gkr1079.
- (35) Rivera-Leon R, Green CJ, Vold BS (1995) High-level expression of soluble recombinant RNase P protein from *Escherichia coli*. *J. Bacteriol.* 177(9):2564-2566.
- (36) Rosenblad MA, Lopez MD, Piccinelli P, Samuelsson T (2006) Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res.* 34(18):5145-5156. doi: 10.1093/nar/gk1626.
- (37) Shigi N, Sakaguchi Y, Suzuki T, Watanabe K (2006) Identification of two tRNA thiolation genes required for cell growth at extremely high temperatures. *J. Biol. Chem.* 281(20):14296-14306. doi: 10.1074/jbc.M511675200.
- (38) Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75.
- (39) Stagno JR, Altieri AS, Bubunenko M, Tarasov SG, Li J, Court DL, Byrd RA, Ji X (2011) Structural basis for RNA recognition by NusB and NusE in the initiation of transcription antitermination. *Nucleic acids research* 39(17):7803-7815, doi:10.1093/nar/gkr418.
- (40) Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41(1):207-234.
- (41) Symmons MF, Jones GH, Luisi BF (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* 8(11):1215-1226.
- (42) Taschner A, Weber C, Buzet A, Hartmann RK, Hartig A, Rossmannith W. (2012) Nuclear RNase P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. *Cell Rep.* Jul 26;2(1):19-25
- (43) Tsai HY, Pulukkunat DK, Woznick WK, Gopalan V. (2006) Functional reconstitution and characterization of *Pyrococcus furiosus* RNase P. *Proc. Natl. Acad. Sci. USA* 103(44):16147-16152.
- (44) Vilardo E, Nachbaur C, Buzet A, Taschner A, Holzmann J, Rossmannith W (2012) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase--extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res.* 40(22):11583-11593 (2012). doi: 10.1093/nar/gks910.

- (45) Walczyk D, Gößringer M, Rossmannith W, Zatsepin TS, Oretskaya TS, Hartmann RK (2016) Analysis of the Cleavage Mechanism by Protein-Only RNase P Using Precursor tRNA Substrates with Modifications at the Cleavage Site. *J. Mol. Biol.* 428(24 Pt B):4917-4928. doi: 10.1016/j.jmb.2016.10.020.
- (46) Wang MJ, Gegenheimer P (1990) Substrate masking: binding of RNA by EGTA-inactivated micrococcal nuclease results in artifactual inhibition of RNA processing reactions. *Nucleic Acids Res.* 18(22):6625-6631.
- (47) Weber C, Hartig A, Hartmann RK, Rossmannith W (2014) Playing RNase P evolution: swapping the RNA catalyst for a protein reveals functional uniformity of highly divergent enzyme forms. *PLoS Genet.* 7;10(8)
- (48) Wegscheid B, Hartmann RK (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. *RNA.* 12(12):2135-2148. doi: 10.1261/rna.188306.
- (49) Xu Y, Amero CD, Pulukkunat DK, Gopalan V, Foster MP. (2009) Solution structure of an archaeal RNase P binary protein complex: formation of the 30-kDa complex between *Pyrococcus furiosus* RPP21 and RPP29 is accompanied by coupled protein folding and highlights critical features for protein-protein and protein-RNA interactions. *J. Mol. Biol.* 393(5):1043-1055. doi: 10.1016/j.jmb.2009.08.068.
- (50) Brillante N, Gößringer M, Lindenhofer D, Toth U, Rossmannith W, Hartmann RK. (2016) Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P. *Nucleic Acids Res.* Mar 18;44(5):2323-36

## 4.2 Publikation 2:

### **Homologs of *Aquifex aeolicus* protein-only RNase P are not the major RNase P activities in the archaea**

#### ***Haloferax volcanii* and *Methanosarcina mazei***

Im Anschluss an die Identifizierung der homologen Proteine von AaRP (HARP, homolog of *A. aeolicus* RNase P) (4.1) wurde die Frage adressiert, welche Funktion die HARP-Proteine in den jeweiligen Organismen besitzen. Um dieser Fragestellung nachzugehen wurden stellvertretend Archaeen ausgewählt, welche gentechnisch verändert werden können. Dabei handelte es sich um *H. volcanii* und *M. mazei*.

Zunächst wurden die rekombinanten HARP Proteine hinsichtlich ihrer katalytischen Aktivität untersucht. Dabei konnte bei dem HARP von *H. volcanii* (Hv-HARP) *in vitro* eine 5'-Prozessierungsaktivität nachgewiesen werden, allerdings zeigte das Protein keine *in vivo*-Funktionalität in *E. coli* (Fig. 4 (Seite 130); Fig. 5 (Seite 131)). Hingegen konnte bei dem HARP von *M. mazei* (Mm-HARP) sowohl *in vitro* als auch *in vivo* eine funktionale RNase P Aktivität beobachtet werden (Fig. 4 (Seite 130); Fig. 5 (Seite 131)). Im Vergleich zu Hv-HARP ist die Prozessierung durch Mm-HARP etwas reduziert, was auf den Analysen-Gelen mit bloßem Auge zu erkennen ist. Es wäre möglich, dass aufgrund der N-terminalen Histidin-Fusion die Bildung des aktiven Zentrums gestört wird. Dies konnte in früheren vergleichenden Experimenten mit *T. indicus*-HARP gezeigt werden. Wie zuvor bei Präparationen von rekombinantem AaRP konnte auch bei denen von Hv-HARP durch Kontrollexperimente eine putative Kontamination durch *E. coli* RNase P ausgeschlossen werden (Fig. S5 (A) (Seite 146)). In Mm-HARP Fraktionen konnte eine sehr niedrige Konzentration an *E. coli* RNase P RNA nachgewiesen werden. Allerdings konnte in einem Kontrollansatz mit *in vitro*-assembliertem *E. coli* Holoenzym bei dieser Konzentration keine Prozessierungsaktivität mehr gemessen werden (Fig. S5 (Seite 146)).

Im Rahmen der Komplementationsanalyse mit dem *E. coli*-Stamm BW konnte gezeigt werden, dass *M. mazei* HARP in der Lage ist den letalen Phänotyp zu komplementieren, im Gegensatz zu *H. volcanii* HARP (Fig. 4 (Seite 130)). Zusätzlich zeigte sich, dass sich das rekombinant exprimierte Hv-HARP



zum großen Teil in der löslichen Zellfraktion befindet (Fig. S4 (Seite 145)). Dadurch konnte ausgeschlossen werden, dass ein Löslichkeitsproblem für die nicht-funktionale Komplementation verantwortlich war.

Nachdem anhand der rekombinanten Proteine eine RNase P-Aktivität bei den HARP-Proteinen von *H. volcanii* und *M. mazei* nachgewiesen werden konnte, wurde sowohl von einem *H. volcanii*- als auch von einem *M. mazei*-Stamm eine entsprechende HARP-Deletionsmutante konstruiert. In beiden Fällen konnte das HARP chromosomal deletiert werden, was durch entsprechende Kontrollen überprüft wurde (Fig. S2 (Seite 143); Fig. S3 (Seite 144)). Wachstumsanalysen zeigten sowohl für *H. volcanii* (Fig. 3 (Seite 129)), als auch für *M. mazei* (Fig. 2 (Seite 128)) keinen veränderten Phänotyp bei einer Deletion des HARPs.

Aus den vorhandenen Daten kann man schließen, dass die klassische RNA-basierte RNase P die Hauptfunktion bei der Katalyse der 5'-Flanken von Prä-tRNAs ausübt. Aussagen zu einer möglichen physiologischen Funktion der HARPs in den Organismen können zum jetzigen Zeitpunkt noch nicht gemacht werden. Allerdings konnte in dieser Arbeit nachweislich gezeigt werden, dass die analysierten rekombinanten HARP-Proteine *in vitro* eine RNase P-Aktivität besitzen.

## 4.2.1 Originalpublikation

## Research Communication



# Homologs of *Aquifex aeolicus* Protein-Only RNase P Are Not the Major RNase P Activities in the Archaea *Haloferax volcanii* and *Methanosarcina mazei*

Thandi S. Schwarz<sup>2†</sup>  
 Nadine B. Wäber<sup>1†</sup>  
 Rebecca Feyh<sup>1</sup>  
 Katrin Weidenbach<sup>3</sup>  
 Ruth A. Schmitz<sup>3</sup>  
 Anita Marchfelder<sup>2\*</sup>   
 Roland K. Hartmann<sup>1\*</sup>

<sup>1</sup>Institute of Pharmaceutical Chemistry, Philipps University, Marburg, Germany

<sup>2</sup>Biology II, Ulm University, Ulm, Germany

<sup>3</sup>Institute of General Microbiology, Christian-Albrechts-Universität, Kiel, Germany

## Abstract

The mature 5'-ends of tRNAs are generated by RNase P in all domains of life. The ancient form of the enzyme is a ribonucleoprotein consisting of a catalytic RNA and one or more protein subunits. However, in the hyperthermophilic bacterium *Aquifex aeolicus* and close relatives, RNase P is a protein-only enzyme consisting of a single type of polypeptide (Aq\_880, ~23 kDa). In many archaea, homologs of Aq\_880 were identified (termed HARP for Homologs of *Aquifex* RNase P) in addition to the RNA-based RNase P, raising the question about the functions of HARP and the classical RNase P in these archaea. Here we investigated HARPs from two euryarchaeotes, *Haloferax volcanii* and *Methanosarcina mazei*. Archaeal strains with HARP gene knock-outs showed no growth phenotypes under standard conditions, temperature and salt stress (*H. volcanii*) or nitrogen deficiency (*M. mazei*). Recombinant *H. volcanii* and *M. mazei* HARPs were basically able to

catalyse specific tRNA 5'-end maturation in vitro. Furthermore, *M. mazei* HARP was able to rescue growth of an *Escherichia coli* RNase P depletion strain with comparable efficiency as Aq\_880, while *H. volcanii* HARP was unable to do so. In conclusion, both archaeal HARPs showed the capacity (in at least one functional assay) to act as RNases P. However, the ease to obtain knockouts of the singular HARP genes and the lack of growth phenotypes upon HARP gene deletion contrasts with the findings that the canonical RNase P RNA gene cannot be deleted in *H. volcanii*, and a knockdown of RNase P RNA in *H. volcanii* results in severe tRNA processing defects. We conclude that archaeal HARPs do not make a major contribution to global tRNA 5'-end maturation in archaea, but may well exert a specialised, yet unknown function in (t)RNA metabolism. © 2019 IUBMB Life, 71(8):1109–1116, 2019

**Keywords:** tRNA processing; 5'-end maturation; RNase P; homologs of *Aquifex* RNase P; HARPs; evolution of RNase P

Additional Supporting Information may be found in the online version of this article.

**Abbreviations:** cHis, C-terminal His tag; CRISPR, clustered regularly interspaced short palindromic repeats; CRISPRi, CRISPR interference; DTT, 1,4-Dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HARP, Homolog of *Aquifex* RNase P; MRPP, mitochondrial RNase P protein; nHis, N-terminal His tag; pac, Puromycin N-acetyltransferase resistance gene cassette; PilT protein, type IV pili twitching motility protein; PIN, PilT N-terminal domain-like; Pop5, processing of precursor 5, RNase P protein subunit; PPR, pentatricopeptide repeat; pre-tRNA<sup>Gly</sup>, precursor tRNA<sup>Gly</sup>; PRORP, protein-only RNase P; RNase P, Ribonuclease P; Rpp, RNase P protein; rRNA, ribosomal RNA; tmRNA, transfer messenger RNA; tRNA, transfer RNA; VapB, virulence-associated protein B; VapC, virulence-associated protein C

© 2019 International Union of Biochemistry and Molecular Biology

Volume 71, Number 8, August 2019, Pages 1109–1116

\*Address correspondence to: Anita Marchfelder, Biology II, Ulm University, Ulm, Germany. E-mail: anita.marchfelder@uni-ulm.de

Roland K. Hartmann, Institute of Pharmaceutical Chemistry, Philipps University, Marburg, Germany. E-mail: roland.hartmann@staff.uni-marburg.de

<sup>†</sup>Thandi S. Schwarz and Nadine B. Wäber share the joint first authorship.

Received 9 May 2019; Accepted 9 June 2019

DOI 10.1002/iub.2122

Published online 8 July 2019 in Wiley Online Library

(wileyonlinelibrary.com)



## INTRODUCTION

In all domains of life (Bacteria, Archaea and Eukarya), the mature 5'-ends of tRNA molecules are generated by the essential endonuclease RNase P. The enzyme was originally identified to consist of a catalytic RNA and a single protein subunit in bacteria such as *Escherichia coli* and *Bacillus subtilis* (1). This RNA-centric architecture was subsequently found in almost all bacterial subphyla as well as in archaea and eukarya, in the two latter phylogenetic domains in the form of a ribonucleoprotein composed of a single RNA and 4–5 (archaea) and 9–10 (eukaryotic nuclei) different protein subunits (2–10). The ribonucleoprotein type of RNase P is assumed to be the ancient form of the enzyme.

Although several reports had provided initial hints on the possible existence of protein-only RNase P activities (11–14), the first unequivocal demonstration of a protein-only RNase P was reported in 2008 by Rossmann and coworkers who identified the components of human mitochondrial RNase P (15). The latter is a complex of three different proteins, of which the subunit mitochondrial RNase P protein 3 (MRPP3) catalyzes tRNA 5'-end maturation. Bioinformatic searches for homologs of MRPP3 rapidly identified homologs in other eukarya, but not in archaea or bacteria. It was then shown for the model plant system *Arabidopsis thaliana* that its RNase P activities in the nucleus, chloroplasts and mitochondria reside in single polypeptides orthologous to human MRPP3, which led to coining of the term PRORP for "proteinaceous or protein-only RNase P" (16). Two PRORPs were also demonstrated to be the sole forms of RNase P in the nucleus and mitochondria of the protist *Trypanosoma brucei* (17). Finally, a comprehensive bioinformatic analysis revealed that PRORP proteins are widespread among eukarya, occurring in four of the five eukaryal supergroups, only lacking in the supergroup Amoebozoa (18). The latter study did not reveal any substantial evidence for the coexistence of RNA-based RNase P and PRORP within the same compartment of a eukaryotic cell; such a constellation is of interest as it mimics the functional redundancy that is expected to have occurred in the evolution of cells that substituted one form of RNase P with the other in one of their compartments (nucleus or organelles).

Recently, a new form of RNA-free RNase P (Aq\_880) was also discovered in a bacterium, the hyperthermophile *Aquifex aeolicus* (19). As the eukaryal PRORPs, Aq\_880 is also a member of the PIN domain-like superfamily of metallo-nucleases, but belonging to a different subgroup (PIN\_5 cluster, VapC structural group) (20). In contrast to the known PRORPs (~60–65 kDa (10, 17, 18, 21)), Aq\_880 is a remarkably small protein (~23 kDa), which shows limited similarity to the metallo-nuclease domain of PRORPs but lacks their pentatricopeptide repeat (PPR) RNA-binding and bipartite zinc-binding domains (19). Homologs of Aq\_880, termed Homologs of *Aquifex* RNase P (HARPs), were identified in some bacteria and many archaea (see alignment in Fig. S1). In a few of these bacteria and in all of those archaea, HARP coexists with the classical RNase P consisting of a catalytic RNA and the corresponding protein subunit(s). If both enzymes

acted as RNases P in these organisms, this would be a so far unprecedented co-occurrence of a protein- and an RNA-based RNase P in the same cellular compartment; such a scenario has not even been observed in Eukarya, where the two RNase P forms are nevertheless more widespread. This prompted us to investigate the function of HARP and the classical RNase P in this group of archaea. We therefore characterised the HARP proteins of selected archaeal species that are amenable to genetic studies, that is, in the halophile *Haloferax volcanii* and in the anaerobic *Methanosarcina mazei* (both euryarchaeotes). Strains with HARP gene knockouts were constructed and tested for growth phenotypes. Furthermore, HARPs from the two organisms were analysed for tRNA 5'-end maturation in vitro and for functional complementation of an *E. coli* RNase P mutant strain.

## MATERIALS AND METHODS

### Growth of Strains Used in This Study

In general, *M. mazei* \*, a derivative of the *M. mazei* wild type strain (DSM3647) showing higher plating efficiency due to several serial passages over solid media, was used as the parental strain (22). *M. mazei* \* and its derivative mutant strain *M. mazei mm\_2077::pac* carrying a knockout of the HARP gene were grown in minimal media in the presence of 150 mM methanol and 40 mM acetate as carbon and energy sources. For nitrogen sufficiency, 10 mM ammonium was added. Under nitrogen limiting conditions (nitrogen deficiency), atmospheric N<sub>2</sub> gas served as the sole nitrogen source. Cultures were grown under anaerobic conditions at 37°C without shaking as described previously (22). Optical density was measured at 600 nm.

*H. volcanii* strains were grown in Hv-YPC medium (23) at 45°C and 200 rpm (Innova 43, New Brunswick). Temperature stress experiments were performed at 30°C for cold stress and 48.5°C for heat stress. High and low salt stress was induced in the presence of 23% and 15% salt, respectively. Strains carrying a plasmid were grown in casamino acids medium (Hv-Ca) or minimal medium (Hv-Min) as described (23; <https://haloarchaea.com/halohandbook/>). *E. coli* strain DH5α was grown aerobically at 37°C under shaking at 200 rpm in LB or 2x YT medium. For the *H. volcanii* growth experiments performed in glass tubes, 4 mL of medium were inoculated from an overnight culture to a starting OD<sub>650</sub> of 0.01 when using rich medium. Strains were grown in triplicates. The glass tubes were incubated at 45°C under constant shaking and the OD<sub>650</sub> was measured at various time points along the growth curve. Average absorption values and the standard deviation of each value were calculated based on the three replicates and plotted.

### Construction of ΔHARP Knockouts in *M. mazei* and *H. volcanii*

The construction and verification of the HARP knockout strains is described in detail in the Supporting Information.

### Cloning of HARP Genes, Recombinant HARP Expression and Purification

Cloning of HARP genes (with and without His tag) into plasmid expression vectors as well as expression and purification of recombinant proteins in *E. coli* is detailed in the Supporting Information.

### Complementation Studies of Aq\_880 and HARPs

Derivatives of plasmid pDG148(S/X) (21) encoding the HARP proteins without and with C-terminal His-tag were transformed into *E. coli* BW (24) for complementation experiments carried out as described previously (19).

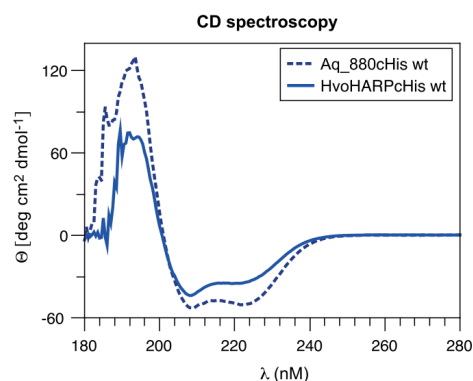
### Pre-tRNA Processing Assays of Aq\_880 and HARPs

The activity of recombinant purified Aq\_880 and archaeal HARPs was analysed essentially as described before (19). Processing assays were carried out in buffer F (50 mM Tris-HCl, pH 7.0, 20 mM NaCl and 4.5 mM MgCl<sub>2</sub>) plus freshly prepared 5 mM DTT; enzyme working solutions were prepared from stock solutions using enzyme dilution buffer (30 mM Tris-HCl, pH 7.8, 30 mM NaCl, 1 mM DTT, 0.3 mM EDTA). For the experiment shown in Fig. 5, 50 nM Aq\_880cHis, 50 nM Hvo\_1114cHis and 500 nM Mm\_2077nHis were incubated with ~5 nM 5'-<sup>32</sup>P-labeled pre-tRNA<sup>Gly</sup> from *Thermus thermophilus* carrying a 14-nt 5'-leader. For more details, see Nickel et al. (19).

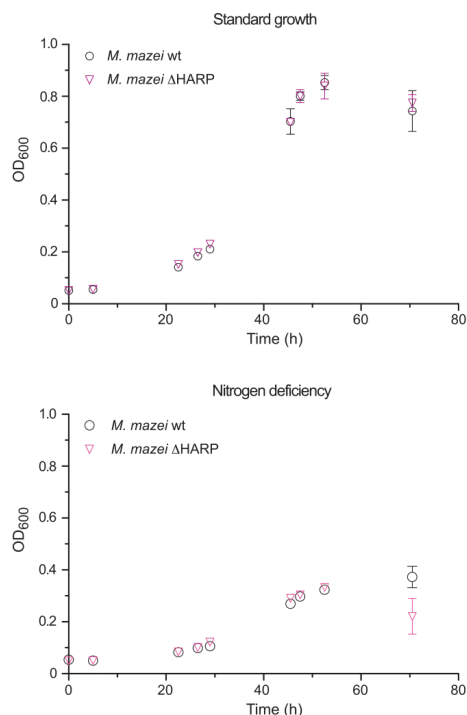
## RESULTS

Homologs of *Aquifex* RNase P (HARPs) can be assigned to the PIN\_5 subgroup within the PIN domain-like superfamily of metallo-nucleases (20); see sequence alignment in Fig. S1). Several Asp residues, shown to be essential for catalysis by Aq\_880

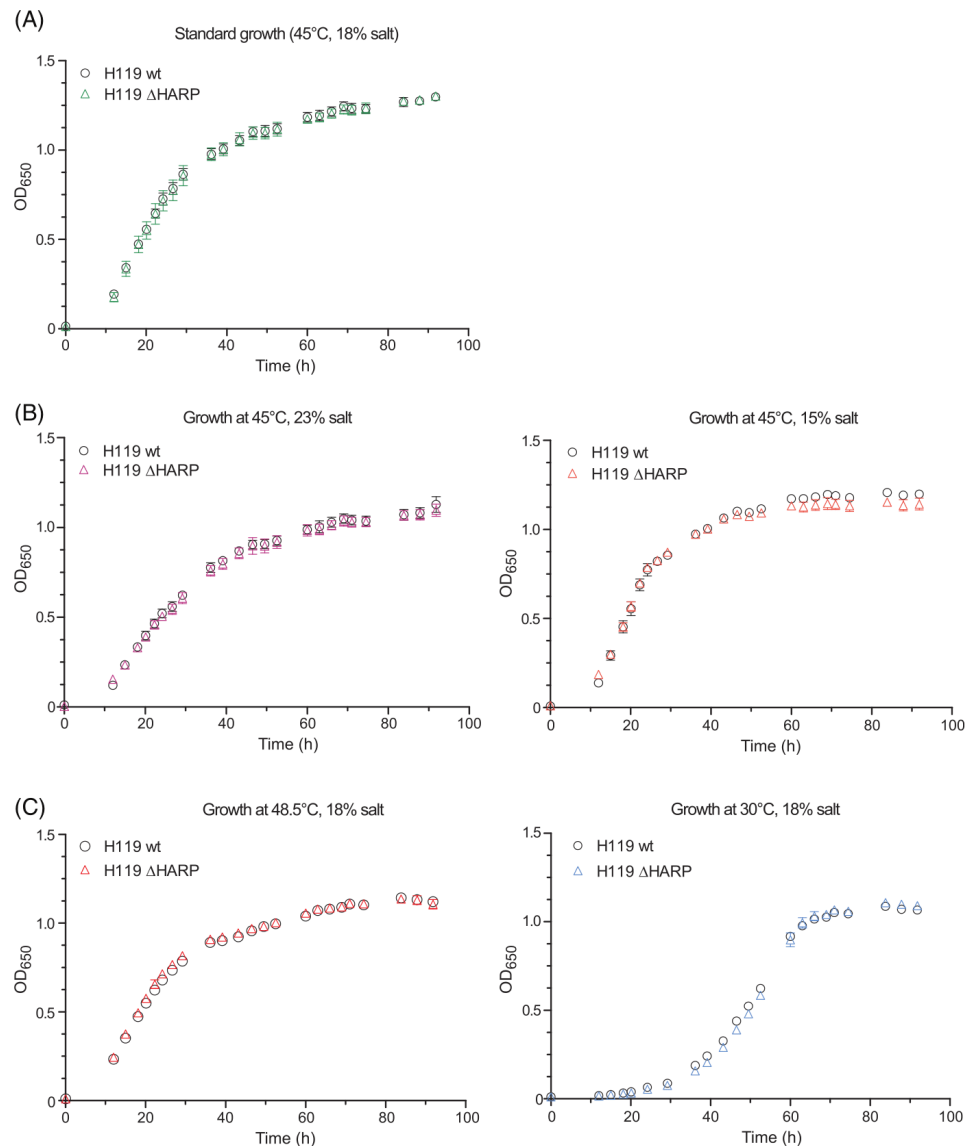
(indicated in Fig. S1), are conserved in this subgroup of metallo-nucleases and contribute to the coordination sphere of one or two catalytic metal ions. CD spectra of recombinant Aq\_880 and *H. volcanii* HARP indicate a predominantly  $\alpha$ -helical fold, which is even more pronounced for the *Aquifex* enzyme (Fig. 1). The native Aq\_880 monomer has a length of 192 aa, while the two archaeal HARPs are longer (232 aa for *H. volcanii* and 247 aa for *M. mazei* HARP; Fig. S1). At present, it is unclear how HARPs specifically recognise pre-tRNA substrates, taking into account that these proteins lack any identifiable domain that may potentially confer RNA binding.



**FIG 1** Circular dichroism (CD) spectra of C-terminally His-tagged Aq\_880 (*Aquifex aeolicus* RNase P) and *H. volcanii* HARP. The spectra indicate that HARPs are predominantly  $\alpha$ -helical.



**FIG 2** Growth curves of the *M. mazei* \* wild type (wt) strain (22) and its derivative mutant strain *M. mazei* mm\_2077::pac carrying a deletion of the HARP gene ( $\Delta$ mm\_2077). Strains were grown in minimal media in the presence of 150 mM methanol and 40 mM acetate as carbon and energy sources. For nitrogen sufficiency, 10 mM ammonium was added. Under nitrogen limiting conditions (nitrogen deficiency), atmospheric N<sub>2</sub> gas served as the sole nitrogen source. Cultures were grown under anaerobic conditions at 37°C without shaking as described previously (22). Optical density was measured at 600 nm.



**FIG 3** Growth curves of *H. volcanii* H119 and its derivative strain with a deletion of the HARP gene ( $\Delta hvo\_1114$ ) under (A) standard conditions in Hv-YPG medium at 45°C, (B) high salt (23%) and low salt (15%) stress, and (C) temperature (48.5°C) and cold stress (30°C). Data points are mean values derived from the three technical replicates; error bars are standard deviations. Optical density was measured at 650 nm.



Here we set out to study the function of archaeal HARPs by pursuing gene knockouts in two archaeal species that are amenable to genetic studies, *H. volcanii* and *M. mazei*. If HARPs played a major role in tRNA 5'-end maturation in these archaeal organisms, one would expect either a failure to generate gene knockouts or at least severe growth defects, taking into account that the function of RNase P in tRNA biogenesis is ubiquitously essential.

#### Growth of Strains with HARP Deletions

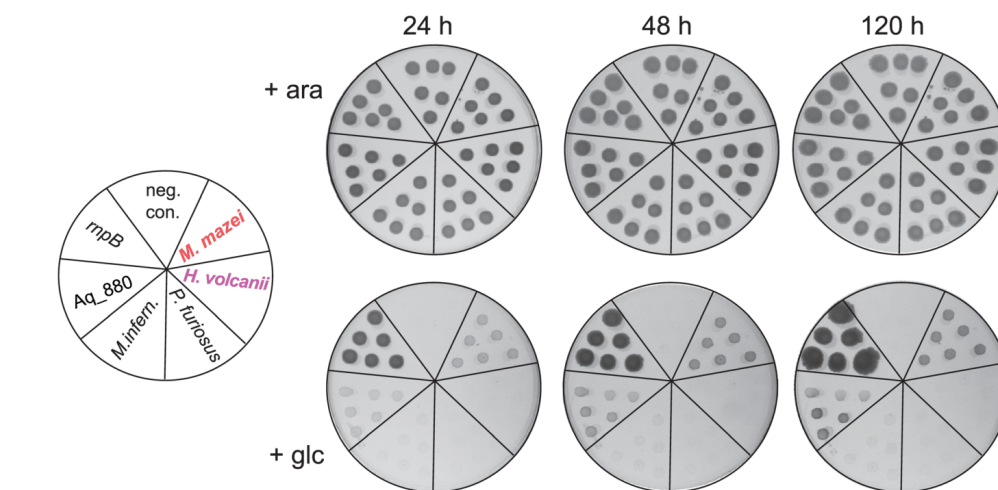
We were able to delete the HARP genes of *M. mazei* (*mm\_2077*) and *H. volcanii* (*hvo\_1114*) as verified by Southern blot analysis (Fig. S2 and S3). Growth of the *M. mazei*  $\Delta$ HARP derivative strain was analysed under standard growth as well as nitrogen deficiency conditions (for details, see Materials and Methods). Under both conditions, growth was essentially indistinguishable from that of the isogenic parental wild type strain (Fig. 2). The relevance of a somewhat lower optical density of the  $\Delta$ HARP strain after 70 h of growth under nitrogen deficiency is considered insignificant, as such drops in optical density after prolonged growth are also seen for wild type *M. mazei* cultures (unpublished observation).

Likewise, growth curves for *H. volcanii* cells lacking the HARP gene were indistinguishable from those of the parental wild type strain (Fig. 3). This was analysed under standard conditions (45°C, 18% salt) as well as high (23%) and low (15%) salt

stress, temperature stress (48.5°C) and cold stress (30°C). A slightly reduced optical density of the  $\Delta$ HARP strain under low salt stress (45°C, 15% salt; Fig. 3B) might suggest a weak phenotype, although this needs to be validated. It is insightful to compare the HARP knockout in *H. volcanii* with corresponding attempts to delete the RNA subunit of the canonical RNase P in the same organism. As such a deletion of the RNase P RNA gene could not be achieved, CRISPRi was used to repress expression of the RNase P RNA subunit. The resulting knockdown (maximally to ~20% of the wild type RNase P RNA level) revealed clear defects in tRNA processing and caused a severe growth defect (25).

#### HARP Complementation of an *E. coli* RNase P Depletion Strain

In the *E. coli* BW strain, chromosomal expression of RNase P RNA (*rnpB*) is switched off when the carbon source is changed from arabinose to glucose (24). We transformed *E. coli* BW cells with plasmids expressing several HARPs, including the ones from *A. aeolicus* (Aq\_880) and *Methylobacterium infernorum* (bacterium) as well as from *M. mazei*, *H. volcanii* and *Pyrococcus furiosus* (archaeon). Plasmid-borne expression of the *E. coli* *rnpB* gene and the plasmid without insert served as positive and negative controls, respectively. As reported previously, Aq\_880 expression rescued growth under non-permissive conditions, although less efficiently than *E. coli*



**FIG 4** Complementation of *E. coli* by HARPs from *M. mazei* and *H. volcanii*. *M. mazei* HARP (*Mm\_2077*), but not *H. volcanii* HARP (*Hvo\_1114*), was able to rescue growth under non-permissive conditions (+ glucose (+ glc)). Rescue strength of *M. mazei* HARP was similar to that of Aq\_880 (RNase P of *Aquifex aeolicus*); *rnpB*: complementation with the *E. coli* RNase P RNA gene as positive control; neg. con.: the empty complementation plasmid used as negative control. *M. infern.*, *P. furiosus*: HARP genes from the bacterium *Methylobacterium infernorum* and the euryarchaeon *Pyrococcus furiosus*. *E. coli* BW cells, transformed with expression plasmids encoding the respective HARPs and controls, were grown at 37°C under permissive (in the presence of arabinose; + ara) or non-permissive (in the presence of glucose; + glc) conditions.

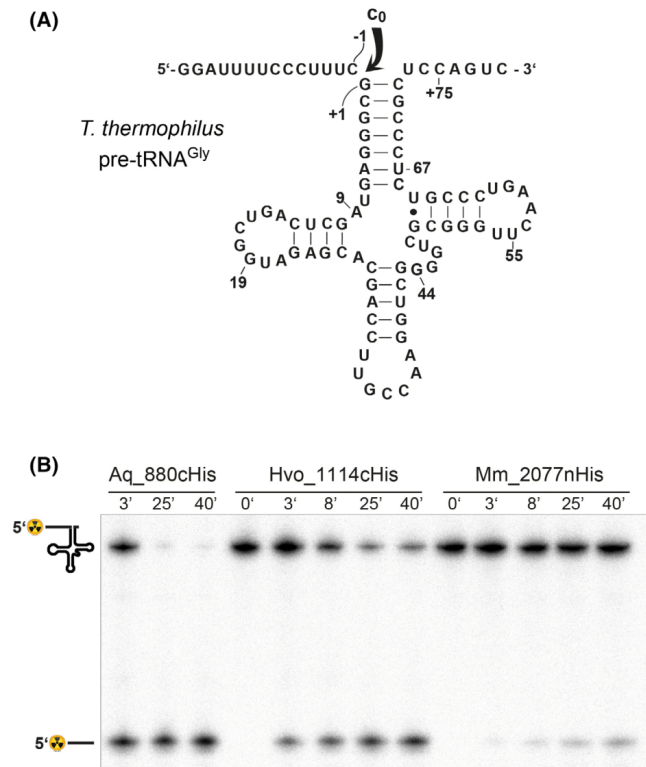


*rnpB* (Fig. 4). Among the other tested HARPs, only the one from *M. mazei* rescued growth of the *E. coli* host to a considerable extent and with similar efficiency as Aq\_880 (Fig. 4). This finding demonstrates that *M. mazei* HARP can functionally replace the endogenous RNA-based *E. coli* RNase P, though with reduced efficacy. In the case of *H. volcanii* HARP, for which RNase P activity was observed in vitro (see later), we analysed if the failure to rescue growth of the RNase P mutant strain may be attributable to insolubility of the halophile protein in the *E. coli* host. However, Western blot analysis revealed that >50% of *H. volcanii* HARP expressed in *E. coli* BW

was in the soluble protein fraction (Fig. S4). This finding does not exclude the possibility that *H. volcanii* HARP adopts a non-native fold in *E. coli*.

#### In Vitro Pre-tRNA Processing by HARPs

Recombinant Hvo\_1114cHis and Mm\_2077nHis were tested for RNase P-specific cleavage under assay conditions that support efficient processing by Aq\_880. Aq\_880cHis, used as control regarding activity and length of the released 5'-flank, was previously demonstrated to cleave the canonical class I precursor tRNA<sup>Gly</sup> (pre-tRNA<sup>Gly</sup>, Fig. 5A) at the canonical site (19). Both



**FIG 5**

(A) Secondary structure of the class I pre-tRNA<sup>Gly</sup> from the thermophilic bacterium *Thermus thermophilus* carrying a 14-nt 5'-leader. The canonical RNase P cleavage site (between nt -1 and +1, termed C<sub>0</sub> site; (26) is indicated. (B) RNase P processing of 5'-<sup>32</sup>P-labeled *T. thermophilus* pre-tRNA<sup>Gly</sup> by recombinant HARPs from the hyperthermophilic bacterium *A. aeolicus* (Aq\_880cHis), the halophilic archaeon *H. volcanii* (Hvo\_1114cHis) and the strictly anaerobic methanogenic archaeon *M. mazei* (Mm\_2077nHis). The first two proteins carried a C-terminal His tag, while the *M. mazei* protein was expressed and purified with an N-terminal His tag. Processing assays were carried out in buffer F (50 mM Tris-HCl, pH 7.0, 20 mM NaCl and 4.5 mM MgCl<sub>2</sub>) plus freshly prepared 5 mM DTT. Reaction mixtures contained ~5 nM substrate and 50 nM Aq\_880cHis, 50 nM Hvo\_1114cHis or 500 nM Mm\_2077nHis. Aliquots were withdrawn at the indicated time points and analysed by 20% denaturing PAGE; 0' indicates incubation without enzyme for the same time period as the last aliquot withdrawn from the sample with enzyme.

1114

ARCHAEAL HOMOLOGS OF *AQUIFEX AEOLICUS* PROTEIN-ONLY RNASE P

archaeal HARP proteins also processed this substrate at the canonical site (Fig. 5B); the efficiency decreased in the order Aq\_880cHis > Hvo\_1114cHis > Mm\_2077nHis under the tested conditions. As we did not individually optimise the assay conditions for Hvo\_1114cHis and Mm\_2077nHis, one cannot exclude that the two may act more efficiently under other conditions. A possible contribution of *E. coli* RNase P contaminations to the activities observed with purified recombinant Hvo\_1114cHis and Mm\_2077nHis preparations could be excluded (see Fig. S5).

## DISCUSSION

Here we have demonstrated that *M. mazei* and *H. volcanii* HARP are able to act as genuine RNases P in vitro, cleaving a canonical precursor tRNA at the canonical cleavage site. While *M. mazei* HARP was less active in vitro than *H. volcanii* HARP under the tested assay conditions, an inverse correlation was seen in complementation experiments using an *E. coli* RNase P depletion strain: *M. mazei* HARP could rescue growth of *E. coli* with similar efficiency as the type enzyme Aq\_880, whereas *H. volcanii* HARP was unable to do so. We conclude from these findings that the two archaeal HARPs are basically capable of acting as RNases P. On the other hand, the ease with which HARP gene deletions could be obtained in the two archaea and the lack of considerable growth defects of the deletion strains suggests that the RNA-based RNase P, expressed in all archaea that encode HARPs (19), is the major archaeal RNase P activity. This notion receives support from our observation that the RNase P RNA subunit gene could not be deleted in *H. volcanii* and repression of the RNase P RNA gene caused severe defects in tRNA biogenesis (25). It is a possibility that archaeal HARPs exert specialised functions in tRNA biogenesis, possibly under extreme conditions not yet identified. RNA-Seq data (27) (Fig. S6) are consistent with relatively low levels of *hvo\_1114* (HARP) transcripts under standard growth conditions. The same pertains to mass spectrometry proteome data, where abundance of Hvo\_1114 was ranked position 1,406 among a total of 1,724 proteins (Jevtic, Stoll, Pfeiffer, Sharma, Urlaub, Marchfelder, Lenz; manuscript in preparation). Transcriptome and proteome data of *M. mazei* indicate co-expression of HARP (Mm\_2077) and RNA-based RNase P subunits (Mm\_1556 = Rpp21, Mm\_2132 = Rpp29, Mm\_2618 = Rpp30, Mm\_2619 = Pop5, Mm\_2467 = L7Ae) (unpublished data).

The HARPs are evolutionarily linked to toxin-antitoxin systems in bacteria and archaea (20), where the toxin proteins are often sequence-specific endoribonucleases that degrade mRNA, rRNA, tmRNA or tRNA to inhibit protein biosynthesis in response to certain stresses (28). For example, proteins with a PIN domain-like fold of the VapC (virulence-associated protein C) type include VapC endonucleases from *Shigella flexneri* and *Salmonella enterica* Typhimurium LT2 that specifically cleave initiator-tRNA<sup>Met</sup> (but not elongator tRNAs including tRNA<sup>Met</sup>) in the anticodon loop at the junction to the 3'-strand of the

anticodon stem (29). This is reminiscent of RNase P cleavage that also occurs at the junction between single-stranded and double-stranded RNA elements. The genome of *Mycobacterium tuberculosis* is estimated to encode ~45 VapBC toxin-antitoxin systems (VapB = antitoxin) (27, 30). One of those toxins, VapC-mt4 was found to selectively target three out of the 45 *M. tuberculosis* tRNAs by cleaving single internal sites within their anticodon loops. Here cleavage and cleavage site selection depended on the presence of nearby recognition sequences, either 5'-ACGC or 5'-ACUGC (30, 31).

We conclude that archaeal HARPs can basically execute the specific RNase P processing reaction, but the classical RNA-based RNase P seems to be the major RNase P tRNA 5'-end maturation activity in archaea. The specific function of archaeal HARPs needs to be elucidated. It is a possibility that archaeal HARPs may act on specific tRNAs, similar to the aforementioned VapC-like tRNA endonucleases. Finally, the basic capacity of archaeal HARPs to function as RNases P makes plausible why the progenitor of the Aquificaceae, after receiving an archaeal HARP by horizontal gene transfer (19), could easily evolve it to an adequate replacement of the ancient RNA-based enzyme.

## ACKNOWLEDGEMENTS

We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG Schm 1052/9-3, DFG Ma1538/21-1 and DFG HA 1672/19-1).

## REFERENCES

- [1] Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849–857.
- [2] Chamberlain, J. R., Lee, Y., Lane, W. S., and Engelke, D. R. (1998) Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes Dev.* 12, 1678–1690.
- [3] Jarrous, N., and Altman, S. (2001) Human ribonuclease P. *Methods Enzymol.* 342, 93–100.
- [4] Hall, T. A., and Brown, J. W. (2002) Archaeal RNase P has multiple protein subunits homologous to eukaryotic nuclear RNase P proteins. *RNA* 8, 296–306.
- [5] Hartmann, E., and Hartmann, R. K. (2003) The enigma of ribonuclease P evolution. *Trends Genet.* 19, 561–569.
- [6] Kouzuma, Y., Mizoguchi, M., Takagi, H., Fukuhara, H., Tsukamoto, M., et al. (2003) Reconstitution of archaeal ribonuclease P from RNA and four protein components. *Biochem. Biophys. Res. Commun.* 306, 666–673.
- [7] Fukuhara, H., Kifusa, M., Watanabe, M., Terada, A., Honda, T., et al. (2006) A fifth protein subunit Ph1496p elevates the optimum temperature for the ribonuclease P activity from *Pyrococcus horikoshii* OT3. *Biochem. Biophys. Res. Commun.* 343, 956–964.
- [8] Cho, I., Lai, L. B., Susanti, D., Mukhopadhyay, B., and Gopalan, V. (2010) Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14573–14578.
- [9] Klemm, B. P., Wu, N., Chen, Y., Liu, X., Kaitany, K. J., et al. (2016) The diversity of ribonuclease P: protein and RNA catalysts with analogous biological functions. *Biomolecules* 6, 27.





- [10] Gobert, A., Bruggeman, M., and Giegé, P. (2019) Involvement of PIN-like domain nucleases in tRNA processing and translation regulation. IUBMB Life. <https://doi.org/10.1002/iub.2062> Epub ahead of print.
- [11] Rossmannith, W., Tullo, A., Potuschak, T., Karwan, R., and Sbis, E. (1995) Human mitochondrial tRNA processing. J. Biol. Chem. 270, 12885–12891.
- [12] Rossmannith, W., and Karwan, R. M. (1998) Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. Biochem. Biophys. Res. Commun. 247, 234–241.
- [13] Thomas, B. C., Gao, L., Stomp, D., Li, X., and Gegenheimer, P. A. (1995) Spinach chloroplast RNase P: a putative protein enzyme. Nucleic Acids Symp. Ser. 33, 95–98.
- [14] Salavati, R., Panigrahi, A. K., and Stuart, K. D. (2001) Mitochondrial ribonuclease P activity of *Trypanosoma brucei*. Mol. Biochem. Parasitol. 115, 109–117.
- [15] Holzmann, J., Frank, P., Löffler, E., Bennett, K. L., Gerner, C., et al. (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135, 462–474.
- [16] Gobert, A., Gutmann, B., Taschner, A., Gößringer, M., Holzmann, J., et al. (2010) A single Arabidopsis organellar protein has RNase P activity. Nat. Struct. Mol. Biol. 17, 740–744.
- [17] Taschner, A., Weber, C., Buzet, A., Hartmann, R. K., Hartig, A., et al. (2012) Nuclear RNase P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. Cell Rep. 2, 19–25.
- [18] Lechner, M., Rossmannith, W., Hartmann, R. K., Tholken, C., Gutmann, B., et al. (2015) Distribution of ribonucleoprotein and protein-only RNase P in Eukarya. Mol. Biol. Evol. 32, 3186–3193.
- [19] Nickel, A. I., Wäber, N. B., Gößringer, M., Lechner, M., Linne, U., et al. (2017) Minimal and RNA-free RNase P in *Aquifex aeolicus*. Proc. Natl. Acad. Sci. U. S. A. 114, 11121–11126.
- [20] Matelska, D., Steczkiewicz, K., and Ginalski, K. (2017) Comprehensive classification of the PIN domain-like superfamily. Nucleic Acids Res. 45, 6995–7020.
- [21] Gößringer, M., Lechner, M., Brillante, N., Weber, C., Rossmannith, W., et al. (2017) Protein-only RNase P function in *Escherichia coli*: viability, processing defects and differences between PRORP isoenzymes. Nucleic Acids Res. 45, 7441–7454.
- [22] Ehlers, C., Weidenbach, K., Veit, K., Deppenmeier, U., Metcalf, W. W., et al. (2005) Development of genetic methods and construction of a chromosomal glnK(1) mutant in *Methanosarcina mazei* strain Gö1. Mol. Genet. Genomics 273, 290–298.
- [23] Allers, T., Ngo, H. P., Mevarech, M., and Lloyd, R. G. (2004) Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. Appl. Environ. Microbiol. 70, 943–953.
- [24] Wegscheid, B., and Hartmann, R. K. (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. RNA 12, 2135–2148.
- [25] Stachler, A. E., and Marchfelder, A. (2016) Gene repression in Haloarchaea using the CRISPR (clustered regularly interspaced short palindromic repeats)-Cas I-B system. J. Biol. Chem. 291, 15226–15242.
- [26] Walczyk, D., Willkomm, D. K., and Hartmann, R. K. (2016) Bacterial type B RNase P: functional characterization of the L5.1-L15.1 tertiary contact and antisense inhibition. RNA 22, 1699–1709.
- [27] Babski, J., Haas, K. A., Näther-Schindler, D., Pfeiffer, F., Förstner, K. U., et al. (2016) Genome-wide identification of transcriptional start sites in the haloarchaeon *Haloferax volcanii* based on differential RNA-Seq (dRNA-Seq). BMC Genomics 17, 629.
- [28] Masuda, H., and Inouye, M. (2017) Toxins of prokaryotic toxin-antitoxin systems with sequence-specific endoribonuclease activity. Toxins (Basel) 9, pii: E140. <https://doi.org/10.3390/toxins9040140>
- [29] Winther, K. S., and Gerdes, K. (2011) Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proc. Natl. Acad. Sci. U. S. A. 108, 7403–7407.
- [30] Cruz, J. W., Sharp, J. D., Hoffer, E. D., Maehigashi, T., Vvedenskaya, I. O., et al. (2015) Growth-regulating *Mycobacterium tuberculosis* VapC-mt4 toxin is an isoacceptor-specific tRNase. Nat. Commun. 6, 7480.
- [31] Sharp, J. D., Cruz, J. W., Raman, S., Inouye, M., Husson, R. N., et al. (2012) Growth and translation inhibition through sequence-specific RNA binding by *Mycobacterium tuberculosis* VapC toxin. J. Biol. Chem. 287, 12835–12847.

## 4.2.2 Supplementary Informationen

### Supporting Information

**Homologs of *Aquifex aeolicus* protein-only RNase P are not the major RNase P activities in the archaea *Haloferax volcanii* and *Methanosarcina mazei***

Thandi S. Schwarz<sup>2,§</sup>, Nadine B. Wäber<sup>1,§</sup>, Rebecca Feyh<sup>1</sup>; Katrin Weidenbach<sup>3</sup>, Ruth A. Schmitz<sup>3</sup>, Anita Marchfelder<sup>2,§</sup>, Roland K. Hartmann<sup>1,§</sup>

<sup>1</sup> *Institute of Pharmaceutical Chemistry, Philipps University, Marburg, Germany*

<sup>2</sup> *Biology II, Ulm university, Ulm, Germany*

<sup>3</sup> *Institute of General Microbiology, Christian-Albrechts-Universität, Kiel, Germany*

§ *Corresponding authors:*

*anita.marchfelder@uni-ulm.de*

*roland.hartmann@staff.uni-marburg.de*

§ *Joint first authorship*

### Supporting Information, Materials and Methods

#### *Plasmids for construction of the $\Delta$ HARP knockout in *M. mazei* \**

Plasmid pRS1277 (Supporting Information, Table S1) was constructed for the deletion of the chromosomal *mm\_2077* gene of *M. mazei* encoding the homolog of *Aquifex* RNase P (HARP). Approx. 1000 bp of up- and downstream sequences of the gene *mm\_2077* were amplified by PCR using the primer pair MM2077\_KpnI / MM2077\_EcoRI\_rev (Supporting Information, Table S2) for the upstream fragment, and primer pair MM2077\_EcoRI / MM2077\_BamHI\_rev for the downstream fragment, as well as chromosomal DNA from *M. mazei* \* as template. The PCR products of the downstream fragment was cloned into pMCL210 yielding pRS1259; subsequently, in front of this insert, the PCR-amplified upstream fragment was introduced into pRS1259, resulting in plasmid pRS1272. Using the EcoRI restriction site for linearization, the *pac* cassette coding for the puromycin resistance gene under control of a constitutive promotor ( $p_{mcr}$ ) was integrated between the up- and downstream fragment in pRS1272 to construct plasmid pRS1277 (Supporting Information, Table S1).

#### *Construction and verification of the chromosomal *M. mazei* mutant*

*M. mazei* \* was transformed with the NotI-linearized plasmid pRS1277 using liposome-mediated transformation as described (Ehlers et al., 2005). For selection, 5  $\mu$ g/ml puromycin was added to minimal medium plates. Single colonies were used to inoculate fresh liquid minimal media containing puromycin. Such cultures were analyzed by Southern blot analysis using probes specific for the ORF of *mm\_2077* and for the *pac*-cassette as described (Weidenbach et al., 2008).

#### *Generation of strain HV69 ( $\Delta$ HARP)*

For the deletion of the *H. volcanii* HARP gene (*hvo\_1114*), strain H119 (Allers et al., 2004) was transformed with the plasmid pTA131-HvoHARPupdo (Supporting Information, Table S1) to mediate integration of the plasmid into the genome. Pop-in candidates were streaked out on agar plates containing 5-FOA to force pop-out of the plasmid backbone. Pop-out candidates were investigated further by Southern blot analysis using 10  $\mu$ g of genomic DNA digested with Sall and separated on a 0.8 % agarose gel. DNA fragments were transferred to a nylon membrane Hybond™-N+ (GE Healthcare) by capillary blotting. For verification of the *H. volcanii* HARP knockout, two probes were used that bind at different positions in or around the altered region. A 452 bp fragment of the region upstream of *H. volcanii* HARP

was amplified using primers iPHaRP1 and HaRP2. A 334 bp fragment of the gene itself was amplified using oligonucleotides 5-BamHI-HaRP and HaRPint\_rev to serve as a gene-specific probe. The PCR fragments were labeled with [ $\alpha$ - $^{32}$ P]-dCTP and the DECAprime II DNA labeling kit (Thermo Fisher Scientific Life Technologies). The labeled products were used as hybridization probes. This unequivocally identified clones carrying the HARP knockout; the new strain was termed HV69.

#### Cloning of HARP genes

Plasmids encoding Aq\_880 have been described previously (Nickel et al., 2017). The genes encoding HARPs of *H. volcanii* (*hvo\_1114*), *M. mazei* (*mm\_2077*), *Pyrococcus furiosus* (*pf\_09185*) and *Methylophilum infernorum* (*minf\_0007*) were synthesized as variants codon-optimized for *E. coli* (except for *minf\_0007*) and cloned into a pUC57 backbone by BIOMATIK (Wilmington, Delaware, USA). Subsequently, the coding portions were excised from the pUC57 backbone using the restriction enzymes NcoI and XhoI, purified by agarose gel electrophoresis and inserted into the NcoI/XhoI sites of pET28a(+). These constructs were transformed into *E. coli* BL21 (DE3) for protein expression. For expression of *M. mazei* HARP carrying an N-terminal His tag, the primer pair 178\_Mm\_pET / 179\_Mm\_pET and plasmid pDG148(+)\_mm\_2077cHis (for construction, see below) as template were used to generate a PCR fragment that was digested with NdeI und Bpu1102I and cloned into pET28a(+) cut with the same enzymes. The resulting plasmid pET28a(+)\_mm\_2077nHis was transformed into the *E. coli* expression strain Lemo21(DE3). For complementation experiments in *E. coli* BW, HARP genes were amplified by PCR as follows:

- for *H. volcanii* HARP without His tag: primer pair 175\_HaVo\_pDG / 176\_HaVo\_pDG rev and the pET28a(+)\_hvo\_1114 vector as template;
- for *H. volcanii* HARP with C-terminal His tag: primer pair 175\_HaVo\_pDG / 646\_cHis\_pDG and the pET28a(+)\_hvo\_1114 vector as template;
- for *P. furiosus* HARP without His tag: primer pair 171\_PyFu\_pDG / 172a\_PyF\_pDG and pET28a(+)\_pf\_09185 as template;
- for *P. furiosus* HARP with C-terminal His tag: primer pair 171\_PyFu\_pDG / 646\_cHis\_pDG and pET28a(+)\_pf\_09185 as template;
- for *M. mazei* HARP without His tag: primer pair 173\_Mm\_pDG / 174\_Mm\_pDG and pET28a(+)\_mm\_2077 as template;
- for *M. mazei* HARP with C-terminal His tag: primer pair 173\_Mm\_pDG / 646\_cHis\_pDG and pET28a(+)\_mm\_2077 as template;
- for *M. infernorum* HARP with C-terminal His tag: primer pair 177a\_Minf\_pDG\_blunt / 646\_cHis\_pDG and pET28a(+)\_minf\_0007 as template.

The aforementioned first six PCR products were digested with SmaI/PaeI, the last construct (*minf\_0007*) only with SmaI, followed by insertion into pDG148(+) cut with the same enzymes. For *M. infernorum* HARP without His tag, the synthetic gene was excised from the pUC57 derivative plasmid using SmaI/PaeI and cloned directly into pDG148(+) via the same restriction sites. pDG derivatives encoding the HARP proteins without and with C-terminal His-tag were transformed into *E. coli* BW (Wegscheid and Hartmann, 2006) for complementation experiments (see below). All plasmids were verified by sequencing (SeqLab, Göttingen, Germany; or MWG Eurofins Genomics, Ebersberg, Germany).

*Expression and purification of recombinant M. mazei (Mm\_2077nHis) and H. volcanii (Hvo\_1114cHis) HARPs*

Expression of Mm\_2077nHis was carried out in *E. coli* Lemo21(DE3). 500 mL LB medium supplemented with 30 µg/ml chloramphenicol and 50 µg/ml kanamycin were inoculated with 5 mL of an overnight culture of Lemo21(DE3) harboring plasmid pET28a(+)-*mm\_2077nHis*. The culture was grown under shaking (200 rpm, GFL 3033 warm air shaker) at 37°C to an OD<sub>600</sub> of 0.3 to 0.8, followed by addition of 0.4 mM IPTG for induction of expression. After shaking the culture for another 6 h at 37°C, cells were harvested by low speed centrifugation. The cell pellet was resuspended in 20 mL NPI-20 buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM imidazole, pH 8.0), followed by addition of lysozyme (f.c. 0.1 mg/mL) and phenylmethylsulfonyl fluoride (PMSF, f.c. 20 µg/mL). Cells were disrupted by sonication on ice (Branson Sonifier 250; 5 x 1 min, intensity 50%, output 5) and the Mm\_2077nHis protein was purified by affinity chromatography using a HisTrap column (1 mL; GE Healthcare) and an ÄKTA FPLC system. After pre-equilibration of the column with NPI-20 buffer, the extract was loaded onto the column with a flow rate of 1 mL/min; the column was washed with NPI-20 buffer until a stable baseline at 280 nm was reached. The protein was eluted by applying a linear gradient from 20 to 500 mM imidazole over 30 column volumes by mixing NPI-20 and NPI-500 buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 500 mM imidazole, pH 8.0) up to 100% NPI-500. Fractions were analyzed on 12% TGX-Stain-Free gels and protein bands were visualized using a ChemiDoc Imaging System (BioRad). Mm\_2077nHis eluted at ≥ 250 mM imidazole. The protein was dialyzed (molecular weight cutoff 12-14 kDa) against storage buffer 1 (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 3 mM MgCl<sub>2</sub>, 1 mM DTT, 20% glycerol) and stored at -20°C.

The expression of Hvo\_1114cHis was carried out in *E. coli* BL21(DE3) at 37°C using auto induction medium [LB medium, 0.05% (w/v) glucose, 0.2% (w/v) lactose, kanamycin (50 µg/mL)]. Cells were grown overnight at 220 rpm and 37°C and harvested by centrifugation on the next day (1250 x g, 4°C, 45 min). The pellet derived from a 500 mL culture was resuspended in 20 mL NPI-20 buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM imidazole).

Cells were lysed by repeated sonification on ice (Branson Sonifier 250; 3 x 4 min; intensity 50%, output 5). The protein lysate was cleared by centrifugation (8200 x g; 4°C, 30 min) followed by filtration (cutoff 0.22 µm) before loading on an equilibrated HisTrap HP column (1 mL, GE Healthcare). The column was washed with NPI-20 buffer until a stable baseline at 280 nm was achieved. For elution of the protein, a linear gradient from 20 to 500 mM imidazole was applied. The protein eluted above 150 mM imidazole. The fractions containing the protein were pooled, dialyzed against MonoQ column start buffer (20 mM Tris-HCl pH 8.0), and loaded onto an anion exchange column (1 mL, MonoQ 5/50 GL, GE Healthcare) pre-equilibrated with 20 mM Tris-HCl, pH 8.0. After sample loading, the column was washed with 20 mM Tris-HCl (pH 8.0) until absorption at 280 nm reached a stable baseline. For elution, a linear salt gradient from 0 to 1 M NaCl in 20 mM Tris-HCl (pH 8.0) was applied over 10 column volumes (10 mL). Fractions containing the protein were dialyzed against storage buffer 2 (10 mM Tris-HCl, pH 8.0, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 50% glycerol and 3 mM freshly added DTT) overnight at 4°C. Hvo\_1114cHis was stored at -20°C.

#### *Analysis of the solubility of HARP proteins expressed in E. coli BW*

To examine the solubility of HARP proteins expressed in *E. coli* BW, two colonies of the type shown in Fig. 4 of the main manuscript, which were grown on complementation plates in the presence of arabinose, were scraped off, resuspended in 1 mL lysis buffer (25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2 mM EDTA, 0.1 % Triton X-100, 0.1 mg/mL lysozyme), transferred into Lysing Matrix B containing tubes (MP Biomedicals, Santa Ana, California, USA) and lysed using a FastPrep®-24 (4 m/s, 20 s, 3x) (MP Biomedicals, Santa Ana, California, USA). The matrix was removed by centrifugation (1000 x g, 5 min) and the supernatant was transferred into a new 2 mL collection tube followed by centrifugation (17000 x g, 4 °C, 60 min) to collect the non-soluble protein fraction. The supernatant containing the soluble protein fraction was transferred into a new collection tube and the non-soluble protein pellet was redissolved in an equal volume of TUS buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 8 M urea). Proteins were separated by electrophoresis using 12 % TGX-Stain-Free gels (BioRad) and visualized using the ChemiDoc Imaging system (activation period 5 min; exposure time 3.7 s in Fig. S5) followed by Western blotting and detection with specific rabbit antibodies against Hvo\_1114cHis and HRP-conjugated anti-rabbit IgG (goat) using the ChemiDoc Imaging system (exposure 8.6 s in Fig. S5).

#### **Supporting Information, References**

Allers, T., Ngo, H. P., Mevarech, M., Lloyd, R. G. (2004) Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl. Environ. Microbiol.* **70**, 943-953.

- Babski, J., Haas, K. A., Näther-Schindler, D., Pfeiffer, F., Förstner, K. U. et al. (2016) Genome-wide identification of transcriptional start sites in the haloarchaeon *Haloferax volcanii* based on differential RNA-Seq (dRNA-Seq). *BMC Genomics* **17**, 629.
- Delmas, S., Shunburne, L., Ngo, H.-P., Allers, T. (2009) Mre11-Rad50 Promotes Rapid Repair of DNA Damage in the Polyploid Archaeon *Haloferax volcanii* by Restraining Homologous Recombination. *PLoS Genet.* **5**, e1000552.
- Ehlers, C., Weidenbach, K., Veit, K., Deppenmeier, U., Metcalf, W. W., Schmitz, R. A. (2005) Development of genetic methods and construction of a chromosomal glnK(1) mutant in *Methanosarcina mazei* strain Gö1. *Mol. Genet. Genomics* **273**, 290–298.
- Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557–580.
- Nakano, Y., Yoshida, Y., Yamashita, Y., Koga, T. (1995) Construction of a series of pACYC-derived plasmid vectors. *Gene* **162**, 157–158.
- Nickel, A. I., Wäber, N. B., Gößbringer, M., Lechner, M., Linne, U. et al. (2017) Minimal and RNA-free RNase P in *Aquifex aeolicus*. *Proc. Natl. Acad. Sci. USA* **114**, 11121–11126.
- Nicol, J. W., Helt, G. A., Blanchard, S. G. Jr., Raja, A., Loraine, A. E. (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730–2731.
- Wegscheid, B., Hartmann, R. K. (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. *RNA* **12**, 2135–2148.
- Weidenbach, K., Ehlers, C., Kock, J., Ehrenreich, A., Schmitz, R. A. (2008) Insights into the NrpR regulon in *Methanosarcina mazei* Gö1. *Arch. Microbiol.* **190**, 319–332.



## Supporting Information, Tables S1 and S2

**Table S1.** Strains used in this study and plasmids relevant to the gene knockouts

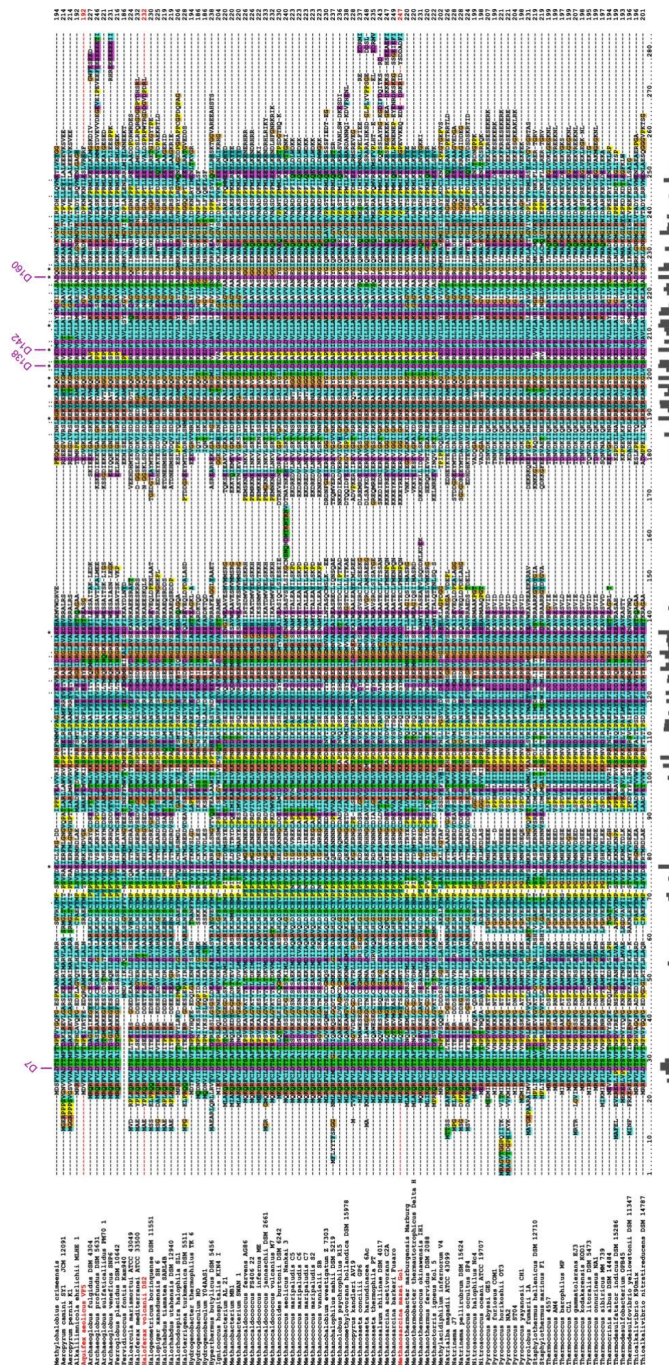
Strains and plasmids	Genotype or description	Source or reference
<b><i>M. mazei</i></b>		
<i>Methanosarcina mazei</i> strain Gö1	wild type	DSM3647
<i>M. mazei</i> *	DSM3647-derived mutant strain with higher plating efficiency	(Ehlers et al., 2005)
<i>M. mazei</i> * MM2077::pac	<i>M. mazei</i> * but <i>mm_2077::pac</i>	this work
<b><i>H. volcanii</i></b>		
H119	DS70(ΔpHV2), ΔpyrE2, ΔtrpA, ΔleuB	(Allers et al., 2004)
HV69	DS70(ΔpHV2), ΔpyrE2, ΔtrpA, ΔleuB, Δhvo_1114	this study
<b><i>E. coli</i></b>		
<i>E. coli</i> DH5α	F <sup>-</sup> , φ80dIacZΔM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(r <sub>k</sub> <sup>-</sup> , m <sub>k</sub> <sup>+</sup> ), phoA, supE44, λ <sup>-</sup> , thi-1, gyrA96, relA1	(Hanahan, 1983), Thermo Fisher Scientific
<i>E. coli</i> BL21AI	F <sup>-</sup> , ompT, gal, dcm, hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> , m <sub>B</sub> <sup>-</sup> ), araB::T7RNAP-tetA, TetA <sup>R</sup>	Thermo Fisher Scientific Invitrogen
<i>E. coli</i> Lemo21(DE3)	<i>fhuA2</i> [lon] <i>ompT gal</i> (λ DE3) [dcm] Δ <i>hsdS</i> / pLemo(CamR) λ DE3 = λ sBamHI ΔEcoRI-B int::( <i>lac</i> ::PlacUV5::T7 gene1) i21 Δ <i>nin5</i> pLemo = pACYC184-PrhaBAD-lysY	New England Biolabs (NEB)
<b>Plasmids (<i>M. mazei</i> genetics)</b>		
pMCL210	general cloning vector	(Nakano et al., 1995)
pRS207	<i>pac</i> -cassette in pBluescript SK <sup>+</sup>	(Ehlers et al., 2005)
pRS1259	<i>mm_2077</i> fragment in pMCL210 (~1000 bp downstream)	this work
pRS1272	pRS1259 and ~1000 bp upstream of ORF <i>mm_2077</i>	this work
pRS1277	pRS1272 with <i>mm_2077::pac</i>	this work
<b>Plasmids (<i>H. volcanii</i> genetics)</b>		
pTA131	ColE1 ori, f1 ori, <i>lacZ</i> , Amp <sup>R</sup> , <i>pyrE2</i>	(Allers et al., 2004)
pTA131-HvoHARPupdo	ColE1 ori, f1 ori, <i>lacZ</i> , Amp <sup>R</sup> , <i>pyrE2</i> , up- and downstream regions of <i>hvo_1114</i>	this study



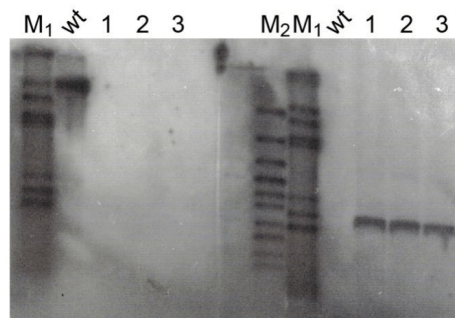
**Table S2.** Oligonucleotides/primers used in this study

Abbreviation	sequence 5'--> 3'	comment
<b>M. mazei genet</b>		
MM2077_KpnI	5'-ATGGGTACCGCATAGTACTCC (KpnI site underlined)	PCR
MM2077_EcoRI_rev	5'-CCGTTGTGTCCAGCAGAATTC (EcoRI site underlined)	
MM2077_EcoRI	5'-TCAGGCATGAATTCGGCTGG (EcoRI site underlined)	
MM2077_BamHI_rev	5'-GGATCCAACATACATGTCCAC (BamHI site underlined)	
<b>H. volcanii genetics</b>		
US	5'-GTAACGCCAGGGTTTTCCC	Colony PCR oligonucleotides
RS	5'-CACAGGAAACAGCTATGAC	
T7	5'-TAATACGACTCACTATAGGG	
T7term	5'-GCTAGTTATTGCTCAGCGG	
HaRP2	5'-ATATCTCGAGGCGATTGACGAACCGAATCACCAGC	Amplification of hvo_1 with flanks up- and downstream
HaRP3	5'-ATATGCGGCCGCGCAAGGCGCGACCGAGCCGATGAC	
iPHaRP1	5'-CGTTCGTCAGTTGGTACCGCG	iPCR for generation of knockout plasmid pTA131-HvoHARPupdo
iPHaRP2	5'-CCCTCCGGTCGCTCTTACC	
HaRPint_rev	5'-GGTTGACCCGGTCGGACA	gene-specific PCR-probe for Southern blot analysis
5S probe	5'-CGCAGGTGAGCTTAACCTCCGTGTTCCGGG	Loading control probe
175_HaVo_pDG	5'-TTTTTCCCGGGATGGCAGAATACCCGCTGAAAC	Hvo_1114 expression plasmid(s)
176_HaVo_pDG_rev	5'-TTTTGCATGCTTACTCGAGACGGTCAGGGTCC	
646_cHis_pDG	5'-AGGCATGCATCCGGATATAGTTCCTCCTTCAG	introduct. cHis tag from pET28a(+)
171_PyFu_pDG	5'-TTTTTCCCGGGATGATCCGTTTCATCCTGGA TACCAGC	Pfc_09185 expression plasmid(s)
172a_PyFu_pDG	5'-TTTTTTGCATGCTTACTTACGTTCTTCTCAC TTTCCAC	
173_Mm_pDG	5'-TTTTCCCGGGATGCTGAAGCAGCGCTTTG	Mm_2077 expression plasmid(s)
174_Mm_pDG	5'-TTTTGCATGCTCACTCGAGAAAATCGGCATCATC GC	
178_Mm_pET	5'-TTTTTTTTTTCATATGCTGAAGCAGCGCTTTGTTCT GG	
179_Mm_pET	5'-TTTTTTTTTGTCTCAGCTCACTCGAGAAAATCGGC ATCATCGC	
177a_MinI_pDG blunt	5'-TTTTTTTCATGAATGAGAAAGTTGTTTTAGA CACCAGC	MinI_0007cHis expression plasmid
<b>E. coli rnpB</b>		
253_EcM1-3	5'-ATCTAGGCCAGCAATCGCTC	RT-PCR of E. coli rnpB (Fig. S5)
254_EcM1-5	5'-GAAGCTGACCAGACAGTCGCCGCTTC	

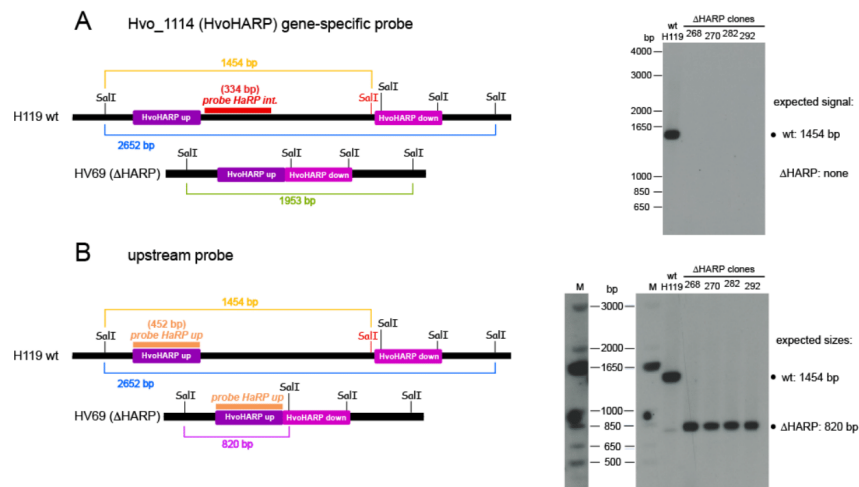
## Supporting Information, Figures S1 to S6



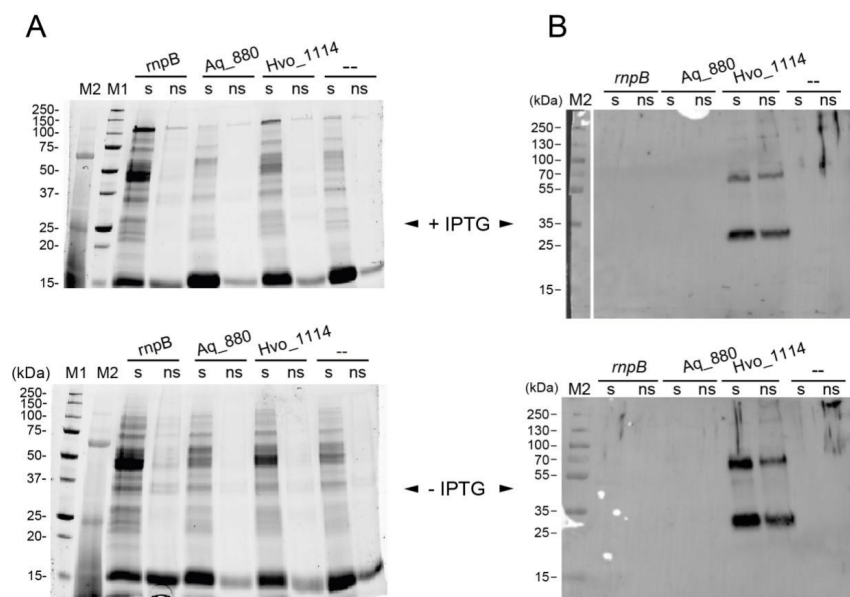
**Fig. S1:** Clustal Omega multiple sequence alignment of Aq\_880 homologs (HARPs) among bacteria and archaea; adapted from Nickel et al. (2017). The sequences of *A. aeolicus* Aq\_880, *H. volcanii* HARP (Hvo\_1114) and *M. mazei* HARP (Mm\_2077) are marked by red lettering on the left and right. The number of amino acids of individual HARPs is indicated in the last column on the right. Above the alignment, the four aspartates shown to be essential for catalytic activity of Aq\_880 (5th sequence from the top) by mutational analysis are indicated in violet (see Nickel et al., 2017, for D138, D142 and D160; this study, unpublished results: D7).



**Fig. S2:** Southern Blot analysis of *M. mazei* \* derivative clones carrying a HARP knockout. Genomic DNA from *M. mazei* \* (wild type, wt) and from 3 independent  $\Delta$ HARP clones (*M. mazei* mm\_2077::pac (1-3) were digested with HindIII and hybridized with a DIG-labeled probe against the ORF mm\_2077 (lanes on the left) and against the pac-cassette (lanes on the right).

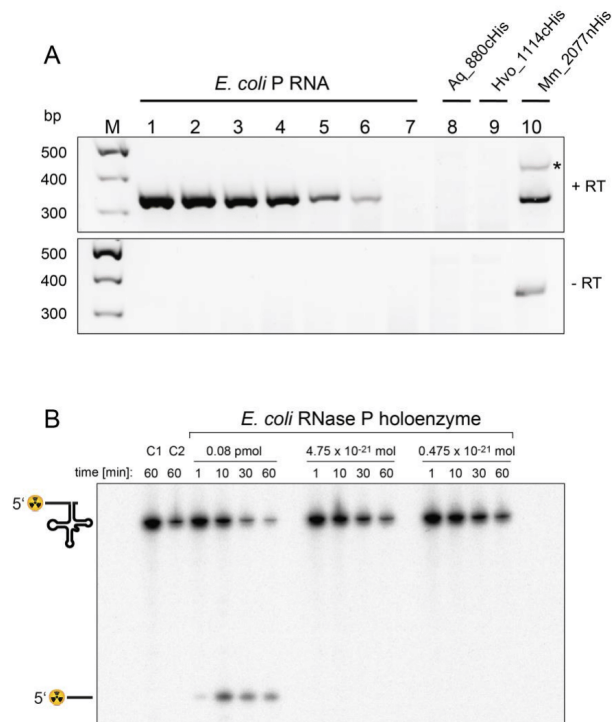


**Fig. S3:** Validation of derivative clones of *H. volcanii* strain H119 carrying a deletion of the HARP gene (*hvo\_1114*). The distance between the two external *SalI* sites is 2652 bp in the H119 wt strain and 1953 bp in the HV69 derivative strain with a deletion of *hvo\_1114* (ΔHARP). Note that the *SalI* site colored in red is deleted in HV69; its distance to the next *SalI* site on the 3'-side is 65 bp in the H119 wt strain. To verify the deletion of *hvo\_1114* (ΔHARP), genomic DNA of several clones was isolated and 10 μg from each clone were digested with *SalI*. DNA fragments were separated on a 0.8% agarose gel, transferred to a nylon membrane (Amersham Hybond™-N<sup>+</sup>, GE Healthcare) and covalently fixed on the membrane by UV crosslinking. **(A)** For hybridization, a 334 bp PCR fragment of the *hvo\_1114* gene (termed probe "HaRP int.") was amplified using primers 5-BamHI-HaRP and HaRPint\_rev; the probe was radioactively labeled with [ $\alpha$ -<sup>32</sup>P]-dCTP using the DECAprime II DNA labeling kit (Thermo Fisher Scientific Life Technologies). **(B)** For detection of genomic DNA upstream of *hvo\_1114*, a 452 bp fragment was amplified using primers iPhaRP1 and HaRRP2; this probe (termed "HaRP up") was labeled in the same manner. Hybridization and detection of membranes were performed according to the manufacturer's protocol. Amersham Hyperfilm ECL films were exposed to membranes for 24 h in **(A)** or 10 h in **(B)**.



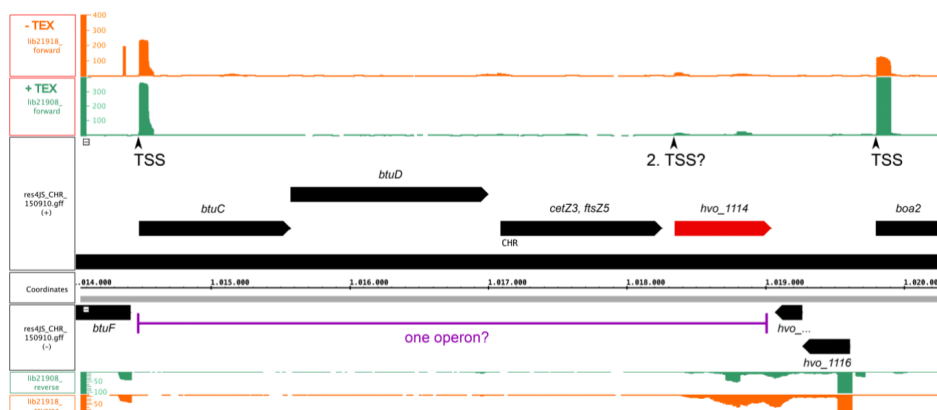
**Fig. S4:** Western blot analysis of the solubility of *H. volcanii* HARP expressed in *E. coli* BW. **(A)** 12% TGX-Stain-Free gels of protein extracts from *E. coli* BW cells grown in the presence of arabinose and transformed with pDG148(+) plasmid vectors expressing *E. coli* RNase P RNA (*rnpB*), Aq\_880 or Hvo\_1114; M1, Precision Plus Protein™ Unstained Protein Standards (Bio-Rad); M2, PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific); sizes of protein marker bands (in kDa) are indicated at the left margin; s, soluble protein fraction; ns, non-soluble protein fraction. **(B)** Western blot analysis of the gels shown on the left using a rabbit antiserum specific for Hvo\_1114cHis. The M1 marker lane was excised in the upper blot image. For further details, see legend to panel A and Experimental Procedures of the main manuscript.





**Fig. S5:** (A) RT-PCR analysis to detect potential *E. coli* P RNA contaminations in preparations of recombinant HARPs. M: DNA marker with fragment sizes (in bp) indicated on the left; Lanes 1 to 7:  $4.75 \cdot 10^{-18}$ ,  $4.75 \cdot 10^{-19}$ ,  $4.75 \cdot 10^{-20}$ ,  $4.75 \cdot 10^{-21}$ ,  $4.75 \cdot 10^{-22}$ ,  $4.75 \cdot 10^{-23}$  and  $4.75 \cdot 10^{-24}$  mol *E. coli* RNase P RNA (*in vitro* transcript), respectively. Lane 8:  $1.25 \cdot 10^{-12}$  mol Aq\_880cHis; lane 9:  $1.25 \cdot 10^{-12}$  mol Hvo\_1114cHis; lane 10:  $12.5 \cdot 10^{-12}$  mol Mm\_2077nHis; the amounts in lanes 8-10 correspond to those used in processing assays (50 nM / 25  $\mu$ L assay volume for Aq\_880cHis and Hvo\_1114cHis, 500 nM / 25  $\mu$ L assay volume for Mm\_2077nHis); the asterisk in lane 10 marks a PCR side product whose appearance depends on reverse transcription. + RT, reverse transcription performed; - RT, reverse transcriptase omitted. The "Access RT-PCR Introductory System" Kit (Promega) was used for RT-PCR with primers 253\_EcM1-3 and 254\_EcM1-5 applying the following temperature steps: step 1: 45 min at 45°C / step 2: 2 min at 94°C / step 3: 35 s at 94°C / step 4: 35 s at 64°C / step 5: 1 min at 68°C / step 6: 7 min at 68°C / step 7: 7°C until termination of PCR; after steps 1 and 2, steps 3-5 were repeated thirty times before proceeding to steps 6 and 7. PCR products were analyzed by electrophoresis on 1.5% agarose gels, stained with GelRed (Biotium) and visualized using the ChemiDoc Imaging System (BioRad). (B) Processing of pre-tRNA<sup>Gly</sup> assayed at different dilutions of *E. coli* RNase P holoenzyme to exclude that traces of *E. coli* P RNA detected in HARP preparations above (panel A) might have contributed to HARP cleavage of pre-tRNA<sup>Gly</sup> in panel A. *E. coli* RNase P holoenzyme was assembled *in vitro* as follows: *in vitro* transcribed *E. coli* P RNA was preincubated in a metal heating block for 1 min at 85°C in 1 x KN buffer [20 mM Hepes-KOH at pH 7.4, 150 mM NH<sub>4</sub>OAc, 2 mM spermidine, 0.05 mM spermine, and 4 mM  $\beta$ -mercaptoethanol] without Mg<sup>2+</sup>; after cooling down to room temperature within ~90 min, Mg(OAc)<sub>2</sub> (f.c. 4.5 nM) and 80 nM recombinant *E. coli* P protein were added, followed by incubation for 5 min at 37°C to assemble functional holoenzymes; holoenzyme preparations were stored at -80°C. For

processing reactions, holoenzyme dilutions (as indicated) were prepared in enzyme dilution buffer (30 mM Tris-HCl pH 7.8, 30 mM NaCl, 1 mM DTT, 0.3 mM EDTA) and preincubated in assay buffer F (50 mM Tris-HCl, pH 7.0, 20 mM NaCl, 4.5 mM MgCl<sub>2</sub>, 5 mM DTT) for 5 min at 37°C; the substrate was preincubated separately for 5 min at 55°C and 5 min at 37°C in the same buffer before combining 16 µL enzyme and 4 µL substrate mix to start the reaction; 4 µL aliquots were withdrawn at indicated time points and mixed with the same volume of ice-cold denaturing gel loading buffer to stop the reaction; samples were analyzed by 20% denaturing PAGE. C1: substrate incubated in buffer F without enzyme; C2: as C1, but in ddH<sub>2</sub>O instead of buffer F.



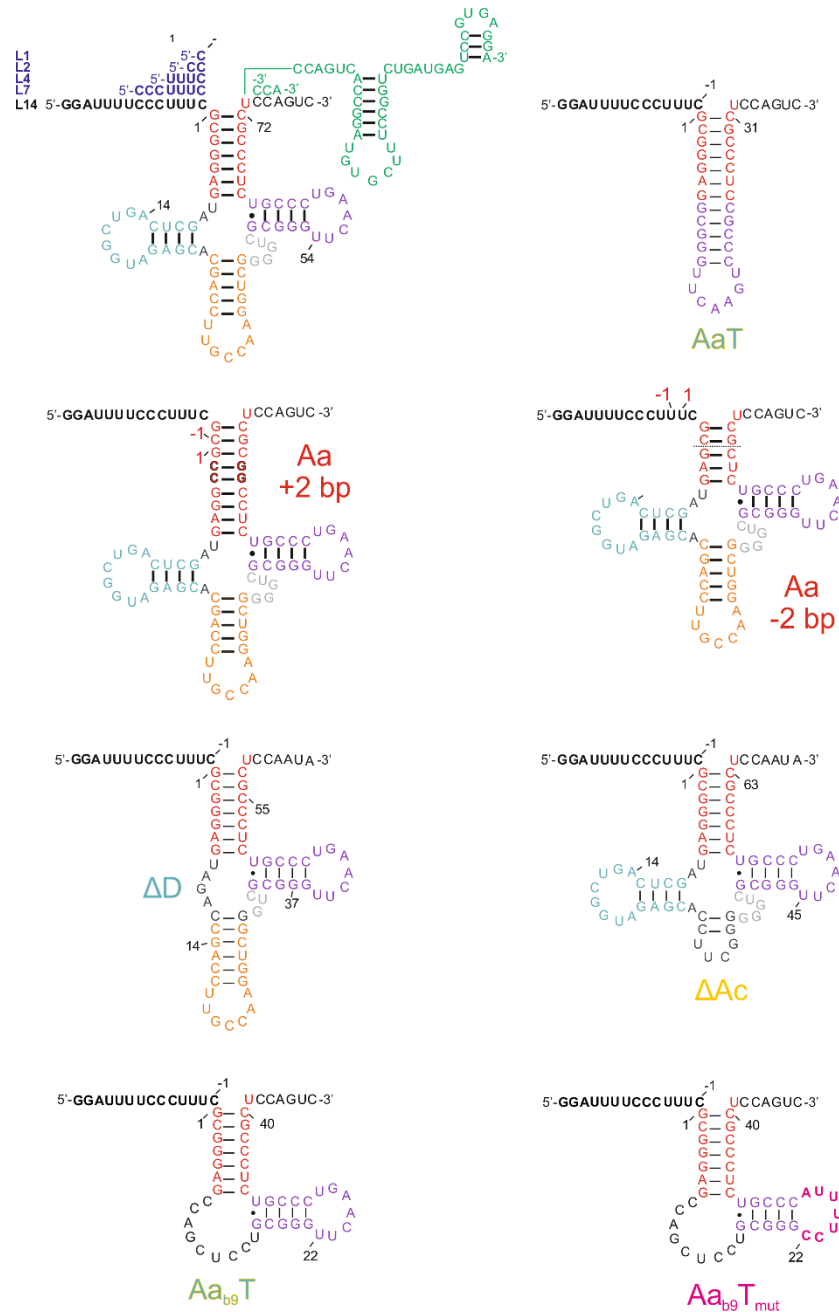
**Fig. S6:** RNA-Seq of total RNA from *H. volcanii* H119 (Babski et al, 2016). Read profiles in an extended region surrounding the HARP gene (Hvo\_1114, red arrow bar) visualized using the Integrated Genome Browser software (Nicol *et al.*, 2009). Reads mapping to the two DNA strands are illustrated at the top and bottom, respectively. Red reads: derived from total RNA without TEX treatment; green reads: pretreatment of total RNA with Terminator™ 5'-Phosphate-Dependent Exonuclease (+ TEX) before cDNA library preparation. A potential transcription start site (TSS) could be identified directly upstream of the *btuB* gene. As the next TSS was only identified at the beginning of the *boa2* gene, it is likely that *btuB*, *btuD*, *cetZ3/ftsZ5* and *hvo\_1114* are part of one operon; however expression of Hvo\_1114 might be driven by a weak promoter immediately upstream of *hvo\_1114* ("2. TSS?"), although the number of reads at this position is very low and there are more (17) reads in the - TEX library than in the + TEX library (12 reads).

### 4.3 Substraterkennung der RNase P von *A. aeolicus*

Wie das neuartige Protein-basierte RNase P-Enzym AaRP seine Substrate erkennt und bindet, war eine weitere Fragestellung dieser Arbeit. Das relativ kleine Enzym (~23 kDa) besteht scheinbar nur aus einer Metallonuklease-Domäne (PIN-5 Gruppe), welche das katalytische Zentrum enthält (NYN-Domäne). Alle bisher bekannten Vertreter der singulären PRORPs sind aus drei Domänen aufgebaut. Sie besitzen eine Metallonuklease-Domäne mit aktivem Zentrum, eine an der Substraterkennung und -bindung beteiligte Pentatricopeptide repeat-Domäne sowie ein zentrales Zink-Finger-Motiv, das die räumliche Anordnung der beiden anderen Domänen zueinander koordiniert (vgl. 1.2.2). Es sollte an Hand von Prozessierungsstudien mit Prä-tRNA-Substratvarianten (Abbildung 5) bestimmt werden, welche Substrateigenschaften für die Erkennung durch AaRP wichtig sind und welche Bereiche von Substrat und Enzym miteinander interagieren. Die analysierten Substratvarianten hatten verschiedene 5'- und 3'-Extensionen, deletierte Bereiche (= Minimalsubstrate), Insertionen und Mutationen.

Die Prozessierung der Substrate mit modifizierten 5'- oder 3'-Extensionen wurde bezogen auf das Substrat mit einem Überschuss an Enzym durchgeführt ( $[E] \gg [S]$ ). Zur Detektion von Substrat und Spaltprodukt wurde das 5'-Ende der Substrate radioaktiv markiert. Nach Zugabe der RNase P-Aktivität wurden nach definierten Zeitpunkten Aliquots aus dem Reaktionsansatz entnommen und auf einem denaturierenden Gel aufgetrennt. Der Reaktionsverlauf konnte durch Auflegen einer Imager-Platte auf das Gel und anschließende densitometrische Auswertung an einem Phosphorimager quantifiziert werden. Auf diese Weise konnten die Parameter  $K_M$  und  $k_{\text{react}}$  bestimmt werden





**Abbildung 5: Strukturvariationen der Prä-tRNA für die Substraterkennung**

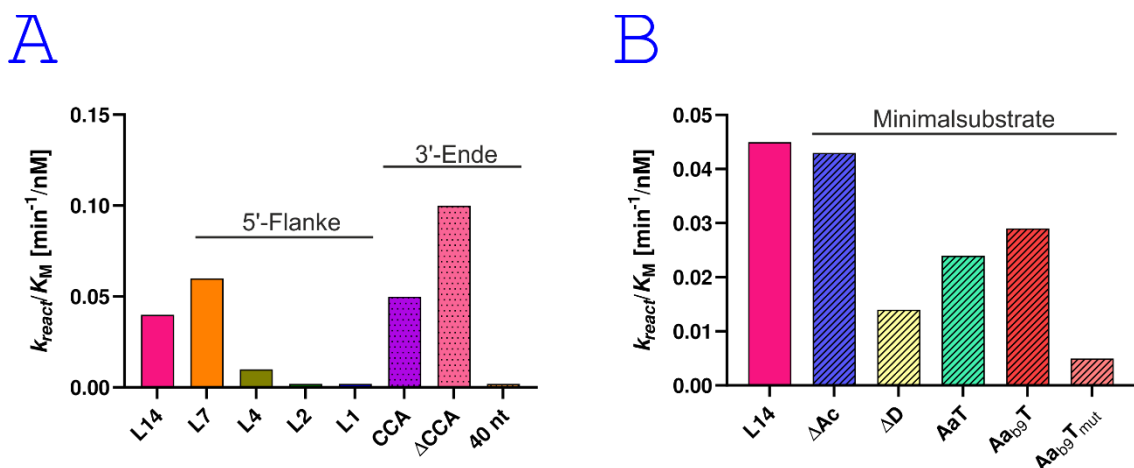
Aufgeführt sind die in dieser Arbeit untersuchten Substratvarianten der Prä-tRNA<sup>Gly</sup> von *Thermus thermophilus* mit einer 5'-Flanke von 14 nt (L14) und einem 3'-Ende von 6 nt (inklusive CCA-Motiv). In der oberen Reihe (links) sind die Variationen in der 5'-Flanke und am 3'-Ende dargestellt sowie rechts das Minimalsubstrat AaT, bei dem die tRNA-Domäne auf den Akzeptorstamm und den koaxial gestapelten T-Arm reduziert ist (rechts). Die zweite Reihe zeigt Varianten der Prä-tRNA<sup>Gly</sup>, bei denen der Akzeptorstamm um 2 Basenpaare verlängert (Aa +2 bp) oder verkürzt (Aa -2 bp) wurde. In den unteren zwei Reihen sind Substrate gezeigt, denen der D-Arm (ΔD) oder der Anticodon Stamm (ΔAc) fehlt sowie zwei vom Minimalsubstrat AaT abgeleitete Varianten, bei denen durch eine Insertion von 9 nt ein leichter Winkel bzw. eine erhöhte Flexibilität zwischen Akzeptorstamm und T-Arm eingeführt wurde; im Substrat Aa<sub>b9</sub>T<sub>mut</sub> wurde zudem die gesamte TΨC-Schleife mutiert, um deren Beitrag zur Enzymbindung zu evaluieren.

#### 4.3.1 Substratvariationen am 5'- oder 3'-Ende der Prä-tRNA

Zunächst wurden die Variationen am 5'- und 3'-Ende untersucht, um festzustellen, ob und welche Interaktionen AaRP mit den Extensionen besitzt. Die Verkürzung der 5'-Flanke der Prä-tRNA<sup>Gly</sup> auf 4 nt (L4) zeigte einen deutlichen Effekt. Der  $K_M$ -Wert erhöhte sich um das 10-fache auf ~300 nM gegenüber Substraten mit 7 nt- bzw. 14 nt-langer 5'-Flanke (L7 bzw. L14). Gleichzeitig konnte eine Verdopplung der Reaktionsgeschwindigkeit auf  $k_{\text{react}} \sim 4 \text{ min}^{-1}$  beobachtet werden. Eine weitere Verkürzung der 5'-Flanke auf eine Länge von nur 1 nt (L1) reduzierte die Spaltungseffizienz  $k_{\text{react}}/K_M$  noch einmal um den Faktor 6,5, wobei sich der  $K_M$ -Wert gegenüber Substrat L4 auf ~630 nM verdoppelte und die Reaktionsgeschwindigkeit von  $4 \text{ min}^{-1}$  auf  $\sim 1 \text{ min}^{-1}$  zurückging (Tabelle 17). Die experimentellen Daten für das Substrat L2 mit einer 2 nt langen 5'-Flanke ergaben eine ungewöhnliche, fast lineare Verteilung im Michaelis-Menten Graph (Abbildung 7B). Entsprechend erhöhte sich der Fehler der Kurvenanpassung für den stark erhöhten  $K_M$ -Wert auf 50% (Tabelle 17). Diese Ergebnisse spiegeln sich auch in der jeweiligen Spaltungseffizienz  $k_{\text{react}}/K_M$  wider. Wie in Abbildung 6A zu erkennen ist, kommt es zu einer starken Verringerung der Spaltungseffizienz bei Verkürzung der 5'-Flanke von 7 nt ( $0,064 \text{ min}^{-1}/\text{nM}^{-1}$ ) auf 4 nt ( $0,013 \text{ min}^{-1}/\text{nM}^{-1}$ ). Bei einer weiteren Verkürzung der 5'-Flanke auf eine Länge von 2 nt bzw. 1 nt können die Substrate zwar noch prozessiert werden, das Verhältnis  $k_{\text{react}}/K_M$  reduziert sich jedoch nochmals auf  $\sim 0,002 \text{ min}^{-1}/\text{nM}^{-1}$ . Diese Ergebnisse weisen auf eine Interaktion zwischen dem Enzym und der 5'-Flanke der Prä-tRNA hin. Für eukaryotische PRORP-Enzyme gibt es dagegen bisher keine experimentellen Hinweise für solch eine Interaktion [107, 108]. Eine Analyse mit dem PRORP3-Isoenzym aus *A. thaliana* zeigte, dass Substrate mit einer Flankenlänge  $\geq 2$  nt alle ähnlich effizient prozessiert werden, lediglich bei einer Länge von 1 nt konnte ein  $k_{\text{react}}$ -Effekt beobachtet werden [108]. Im Gegensatz dazu kommt es bei bakteriellen, RNA-basierten RNase P Enzymen zu einer Interaktion zwischen der Proteinuntereinheit (P Protein) und der 5'-Flanke der Prä-tRNA. Biochemische und strukturelle Studien weisen darauf hin, dass die Nukleotide N-2 bis N-8 der ungepaarten Flanke von der Proteinuntereinheit gebunden werden, wobei die Interaktion mit den Positionen N-4 und N-5 besonders wichtig ist [93, 94, 125, 126]. Dies führt zu einer

Erhöhung der Substrataffinität und –spezifität [66-68, 94, 96]. Folglich wäre bei AaRP ebenfalls denkbar, dass durch die erhöhte Substrataffinität bei der Interaktion mit längeren 5'-Flanken eine bessere Differenzierung zwischen Prä-tRNA und maturer tRNA ermöglicht wird, so dass eine Produktinhibition verhindert wird.

Veränderungen des 3'-Endes der Prä-tRNA zeigten keine nennenswerten Effekte. Das Substrat mit maturem (CCA) und deletiertem ( $\Delta$ CCA) 3'-CCA-Ende wurde mit vergleichbarer Spaltungseffizienz wie das Referenzsubstrat L14 prozessiert, das zusätzlich einen 3 nt-Überhang am CCA-Motiv besitzt. Das Substrat mit einem 40 nt langem Überhang am CCA-Motiv konnte ebenfalls noch gut von AaRP prozessiert werden, allerdings war die Spaltungseffizienz etwas reduziert (Abbildung 6A). Insgesamt konnte eine schwache Präferenz für kürzere 3'-Extensionen beobachtet werden, was jedoch sterische Ursachen haben könnte. Auf Grundlage dieser Ergebnisse können essenzielle Interaktionen zwischen dem 3'-Ende der Prä-tRNA und AaRP ausgeschlossen werden. Während Basenpaarungen zwischen dem CCA-Motiv und der RNA-Untereinheit beim bakteriellen Enzym fundamental wichtig sind [36, 101, 113, 127], spielen vergleichbare Interaktionen bei PRORP offensichtlich keine Rolle bei der Substraterkennung [107, 108]. Folglich ähnelt die Substratbindung an AaRP der von bakterieller RNP RNase P in Bezug auf die Rolle der 5'-Extensionen, während die relative Bedeutungslosigkeit von 3'-Extensionen eine Gemeinsamkeit von AaRP und PRORP-Enzymen ist.



**Abbildung 6: Graphische Darstellung der Spaltungseffizienzen**

(A) Substrate mit variablen 5'-Flanken- und 3'-Enden. L14 entspricht dem Referenzsubstrat (Prä-tRNA<sup>Gly</sup> von *T. thermophilus*). (B) Substrate mit Deletionen (Minimalsubstrate) im Vergleich zum Referenzsubstrat L14.

Tabelle 17: Ergebnisse der Analysen zur Substraterkennung

Es ist das Substrat aufgelistet, inklusive der Eingruppierung, sowie die kinetisch bestimmten Parameter  $k_{\text{react}}$  [ $\text{min}^{-1}$ ],  $K_{\text{M(Sto)}}$  [nM] und die Effizienz  $k_{\text{react}}/K_{\text{M(Sto)}}$  [ $\text{min}^{-1}/\text{nM}^{-1}$ ]

	Substrat	$k_{\text{react}}$ [ $\text{min}^{-1}$ ]	$K_{\text{M(Sto)}}$ [nM]	$k_{\text{react}}/K_{\text{M(Sto)}}$ [ $\text{min}^{-1}/\text{nM}^{-1}$ ]
<b>Referenz</b>	L14	$1,36 \pm 0,15$	$30,3 \pm 8,6$	0,045
<b>5' -Flanke</b>	L7	$1,81 \pm 0,15$	$28,1 \pm 7,3$	0,064
	L4	$4,06 \pm 0,26$	$314,4 \pm 44,4$	0,013
	L2	$4,57 \pm 1,83$	$2221,7 \pm 1109,9$	0,002
	L1	$0,98 \pm 0,11$	$626,8 \pm 116,7$	0,002
<b>3' -Ende</b>	CCA	$5,11 \pm 0,24$	$94,4 \pm 14,6$	0,054
	$\Delta\text{CCA}$	$6,01 \pm 0,30$	$69,9 \pm 10,7$	0,086
	40 nt	$3,65 \pm 0,27$	$127,0 \pm 29,6$	0,029
<b>Minimal Substrate</b>	$\Delta\text{AC}$	$0,68 \pm 0,07$	$15,9 \pm 8,8$	0,043
	$\Delta\text{D}$	$2,67 \pm 0,13$	$192,8 \pm 23,9$	0,014
	AaT	$0,61 \pm 0,06$	$25,4 \pm 11,3$	0,024
	Aa <sub>b9</sub> T	$3,60 \pm 0,21$	$123,9 \pm 17,3$	0,029
	Aa <sub>b9</sub> T <sub>mut</sub>	$0,29 \pm 0,02$	$56,6 \pm 15,0$	0,005

#### 4.3.2 Minimalsubstrate

Die Prozessierung der Minimalsubstrate (Abbildung 5) wurde mit einem Überschuss an Enzym im Vergleich zum Substrat durchgeführt ( $[E] \gg [S]$ ).

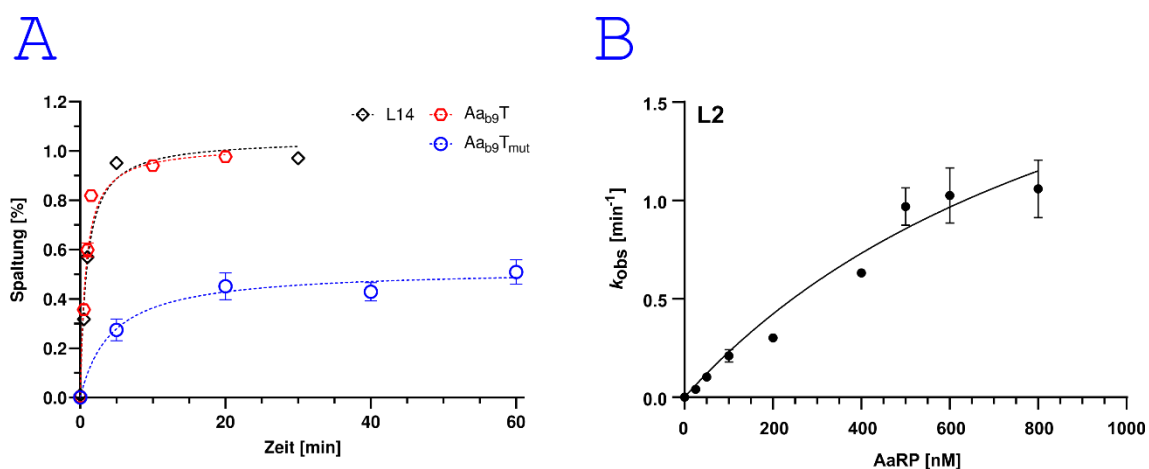
Die Deletion des Anticodon-Arms bei Substrat  $\Delta\text{Ac}$  und die kombinierte Deletion von Anticodon- und D-Arm bei dem Haarnadel-Substrat AaT hatte keinen signifikanten Einfluss auf die Prozessierung durch AaRP. So war die Spaltungseffizienz von  $\Delta\text{Ac}$  ( $0,043 \text{ min}^{-1}/\text{nM}^{-1}$ ) und AaT ( $0,024 \text{ min}^{-1}/\text{nM}^{-1}$ ) vergleichbar mit der des Referenzsubstrats L14 ( $0,045 \text{ min}^{-1}/\text{nM}^{-1}$ ) (Tabelle 17 und Abbildung 6B). Im Vergleich zu Substrat AaT, welches einer rigiden coaxial gestapelten Helix aus Akzeptor- und T-Stamm entspricht, wurden in Substrat Aa<sub>b9</sub>T zusätzlich neun Nukleotide in die 5'-Hälfte zwischen dem theoretischen

Übergang von Akzeptor- und T-Stamm insertiert. Diese Sequenz-Ausbuchtung sollte eine gewisse strukturelle Flexibilität garantieren. Die Konstanten  $K_M$  und  $k_{\text{react}}$  von Spaltungsreaktionen mit Substrat Aa<sub>b9</sub>T waren gegenüber denen von Substrat AaT / L14 zwar leicht erhöht, insgesamt wurde die Spaltungseffizienz dadurch jedoch nicht wesentlich verändert. Die Entfernung des D-Arms hatte im Vergleich zu L14 eine Reduktion der Substrataffinität zur Folge (~6-facher Anstieg von  $K_M$ ), gleichzeitig kam es dadurch aber auch zu einer leichten Erhöhung der Reaktionsgeschwindigkeit (~2-facher Anstieg von  $k_{\text{react}}$ ). Insgesamt führte die Deletion des D-Arms zu einer etwa 3,2-fach verringerten Spaltungseffizienz gegenüber dem Referenzsubstrat L14.

Um mögliche Interaktionen zwischen der TΨC-Schleife der Prä-tRNA und AaRP zu untersuchen, wurde die Sequenz der Schleife bei Substrat Aa<sub>b9</sub>T<sub>mut</sub> vollständig mutiert (Abbildung 5). Mit diesem Substrat wurde der vergleichsweise geringste  $k_{\text{react}}$ -Wert von 0,29 min<sup>-1</sup> gemessen. Selbst bei höheren Enzymkonzentrationen konnten maximal 80% des Aa<sub>b9</sub>T<sub>mut</sub>-Substrats umgesetzt werden. Vergleicht man den  $k_{\text{react}}$ -Wert von 0,29 min<sup>-1</sup> mit dem für Substrat Aa<sub>b9</sub>T (3,6 min<sup>-1</sup>), so wird deutlich, dass die T-Schleife eine wichtige Rolle für die Positionierung des Substrats im aktiven Zentrum spielt.

Die bakterielle RNA-basierte RNase P bindet die typisch L-förmige tRNA hauptsächlich durch Interaktionen mit Akzeptorstamm und T-Arm. Die T-Schleife wiederum wird durch intramolekulare Wechselwirkung mit der D-Schleife in ihrer Position und Konformation stabilisiert. Dennoch können Substrate mit deletiertem D-Arm relativ effizient durch die RNA-basierte RNase P prozessiert werden [100, 103, 128]. Selbst sogenannte Haarnadel-Substrate sind noch gute Substrate der bakteriellen RNase P und werden an spezifischer Position gespalten [97, 129]. Im Gegensatz dazu werden solche Minimalsubstrate nur sehr ineffizient von den eukaryotischen PRORP-Enzymen prozessiert, die nahezu keine Kontakte zu Substratdeterminanten an der Spaltstelle nutzen und z.B. auch nicht mit der 5'-Flanke interagieren. Folglich sind bei den PRORP-Enzymen zusätzliche Wechselwirkungen mit tRNA-Kernstrukturen bedeutend wichtiger als bei den bakteriellen RNA-Enzymen [50, 107, 108]. Auch wenn das Fehlen eines Anticodon-Stamms keinen Effekt auf die Prozessierungseffizienz von PRORP zeigte, so hatte die Deletion des D-Arms einen stark negativen Effekt auf die Substratbindung und -spaltung [107, 108]. Das Haarnadel-Substrat AaT wurde

sehr ineffizient von PRORP aus *A. thaliana* (Isoenzym PRORP3) prozessiert, doch durch Insertion einer 9 nt langen Sequenz (Substrat Aa<sub>b9</sub>T) konnte die strukturelle Flexibilität erhöht und die Spaltungseffizienz deutlich verbessert werden [108]. Bei der Prozessierung durch AaRP hatte die Deletion des D-Arms keinen entscheidenden Einfluss auf die Spaltungseffizienz. Auch das Haarnadel-Substrat AaT wurde noch effizient prozessiert. Hinsichtlich der Substraterkennung der Minimalsubstrate ähnelt AaRP demnach mehr den RNA-basierten bakteriellen Enzymen als den Protein-basierten PRORP-Enzymen. Eine der wenigen konservierten Sequenzelemente der Prä-tRNA befindet sich in der TΨC-Schleife, deren Nukleotide an tertiären Interaktionen mit Nukleotiden in der D-Schleife beteiligt sind [130, 131]. Die Mutation aller Positionen in der Sequenz hatte extreme Effekte auf die Prozessierung durch AaRP und deutet – wie bei den bakteriellen RNase P Enzymen - auf sequenzspezifische Interaktionen hin [97-99]. Allerdings kann nicht ausgeschlossen werden, dass ein durch die Mutationen verursachter struktureller Defekt der TΨC-Schleife die Prozessierung (zusätzlich) beeinflusst hat.



**Abbildung 7: Prozessierung von Prä-tRNA Substratvarianten**

(A) Primärdiagramm zur Bestimmung der Ratenkonstante  $k_{obs}$ . Sowohl mit L14 (◈), als auch mit Aa<sub>b9</sub>T (◈) konnte ein Substratumsatz von fast 100% erreicht werden. Aa<sub>b9</sub>T<sub>mut</sub> (◈) zeigt bei identischer Enzymkonzentrationen einen maximalen Umsatz von 50% des Substrates (AaRP = 50 nM, Fehlerbalken: SEM). (B) Sekundärdiagramm zur Bestimmung von  $k_{react}$  und  $K_{M(sto)}$  von L2. Aufgetragen sind die Ratenkonstanten  $k_{obs}$  [min<sup>-1</sup>] über die Substratkonzentration von AaRP [nM]. Mittels einer an *single turnover*-Bedingungen angepassten Michaelis-Menten-Gleichung [132] wurde eine nicht-lineare Regressionsanalyse der Daten zur Kurvenanpassung durchgeführt ( $n \geq 3$ ; SEM).

### 4.3.3 Akzeptorstamm Variationen

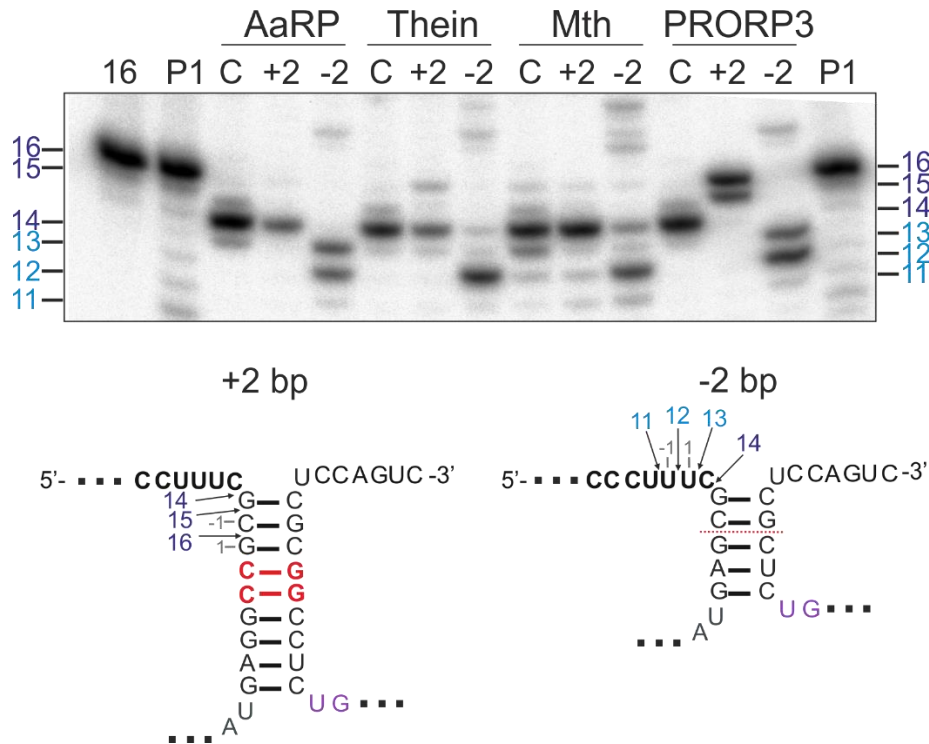
Neben den Variationen der 5'-Flanke, des 3'-Endes und der tRNA-Struktur wurden auch Substrate mit variabler Akzeptorstamm-Länge untersucht. Bei Substrat Aa +2 bp wurden zwei zusätzliche GC-Basenpaare in den Akzeptorstamm inseriert und bei Substrat Aa -2 bp wurden zwei GC-Basenpaare aus dem Akzeptorstamm eliminiert. Die resultierenden Substratvarianten der Prä-tRNA<sup>Gly</sup> besaßen demnach Akzeptorstämme mit 9 bzw. 5 Basenpaaren (Abbildung 5). Anhand dieser Substrate wurde herausgefunden, dass bei PRORP-Enzymen die Spaltstellenselektion über einen „Messmechanismus“ erfolgt. So verschob sich die Spaltstelle entsprechend der Insertion/Deletion um 2 nt in den Akzeptorstamm bzw. in die 5'-Flanke [108]. Bei AaRP hingegen wurde ein anderes Spaltmuster beobachtet. Das Substrat Aa +2 bp wurde von AaRP direkt am Übergang von Akzeptorstamm zu ungepaarter 5'-Flanke gespalten (N<sub>-3</sub>/N<sub>-2</sub>), das heißt, die Spaltstelle wurde um zwei Positionen in 5'-Richtung verschoben (Abbildung 8). Demnach ist AaRP nicht dazu in der Lage, die Akzeptorstamm-Helix vorübergehend aufzuschmelzen, um das Substrat an der kanonischen Spaltstelle zu prozessieren. Die Spaltung des Substrats Aa -2 bp mit verkürztem Akzeptorstamm hingegen erfolgte – ähnlich wie bei PRORP [108] - zu gleichen Teilen an der kanonischen Position N<sub>-1</sub>/N<sub>+1</sub> und an der benachbarten Phosphodiesterbindung N<sub>+1</sub>/N<sub>+2</sub> in der ungepaarten 5'-Flanke. Dies deutet darauf hin, dass auch AaRP eine Art „Messmechanismus“ besitzt, jedoch nicht in der Lage ist, die stabile Akzeptorstamm-Helix partiell zu entwinden oder eine durch thermische Fluktuationen transient zugängliche Spaltstelle gezielt im aktiven Zentrum zu verankern. Außerdem scheint AaRP bevorzugt am Übergang von einer Helix zum Einzelstrang zu schneiden. Diese Tendenz ist auch bei dem Substrat Aa -2 bp zu erkennen, wo die Spaltung zu etwa 50% in Richtung Helix-Ende hin verschoben ist. Insgesamt deutet das Spaltungsmuster von Substrat Aa +2 bp und Aa -2 bp auf eine hohe Flexibilität von AaRP bezüglich der Spaltstellenselektion hin. Zusätzlich wurden die Substrate mit verlängertem bzw. verkürztem Akzeptorstamm auch in Reaktionen mit den HARP-Enzymen des Archaeons *M. thermautotrophicus* und des Bakteriums *T. indicus* getestet. Die Ergebnisse der Spaltstellenanalyse mit den beiden HARP-Enzymen sind ähnlich

zu denen von AaRP. Beide HARP Enzyme spalten Aa +2 bp am Übergang vom gepaartem zum ungepaarten Bereich (Abbildung 8). Während AaRP bei Substrat Aa -2 bp durch die Spaltung an zwei direkt benachbarten Positionen ( $N_{-1}/N_{+1}$  und  $N_{+1}/N_{+2}$ ) eine Doppelbande erzeugt, schneidet *T. indicus* HARP primär an der kanonischen Schnittstelle  $N_{-1}/N_{+1}$  in der 5'-Flanke und bedeutend schwächer am Ende der Akzeptorstamm-Helix an  $N_{-2}/N_{-3}$  in der 5'-Flanke. Bei der Prozessierung mit HARP von *M. thermautotrophicus* konnten deutlich mehr Banden bei dem Substrat Aa -2 bp erkannt werden, primär erfolgte die Spaltung aber wie mit *T. indicus* HARP an den Positionen  $N_{-1}/N_{+1}$  und  $N_{+2}/N_{+3}$ . Demnach erfolgt die Spaltstellenselektion bei HARP-Enzymen ebenfalls mittels einer Art „Längenmessung“ von Substratelementen. Scheinbar besitzen die HARP-Enzyme jedoch eine etwas höhere Flexibilität als AaRP, was es ihnen ermöglicht, auch am Übergangsbereich von Doppel- zu Einzelstrang bei einem um 2 Bp verkürzten Akzeptorstamm zu schneiden. Nicht auszuschließen ist die Möglichkeit, dass die beiden HARPs dafür eine Interaktion mit dem Diskriminatornukleotid eingehen. Eine solche Interaktion ist für die bakterielle RNase P bekannt. So wurde gezeigt, dass das Nukleotid U294 der *E. coli* RNase P RNA die Diskriminator-Base +73 der (Prä)-tRNA kontaktiert [101] um in Kombination mit weiteren Interaktionen die Spaltstelle zu identifizieren.

#### 4.4 Bakterielle Protein-basierte RNase P Enzyme

Durch bioinformatische Suchen konnten homologe Enzyme von AaRP in verschiedenen Bakterien nachgewiesen werden (Publikation I, 4.1). Erstaunlicherweise tragen diese Bakterien, im Gegensatz zu *A. aeolicus*, noch die Gene für die Untereinheiten der RNA-basierten RNase P (*rnpA* und *rnpB*) in ihren Genomen. Zur funktionellen Analyse der HARP-Enzyme wurden repräsentativ die Bakterien *Halorhodospira halophila* (Proteobacteria, halophil) [133], *Alkalilimnicola ehrlichii* (Proteobacteria, halo-alkaliphil) [134], *Thioalkalivibrio nitratireducens* (Proteobacteria, alkaliphil) [135], *Methylophilum infernum* (Verrucomicrobia, acidophil und methanotroph) [136] und *Thermodesulfatator indicus* (Thermodesulfobacteria, thermophil) [137] ausgewählt.



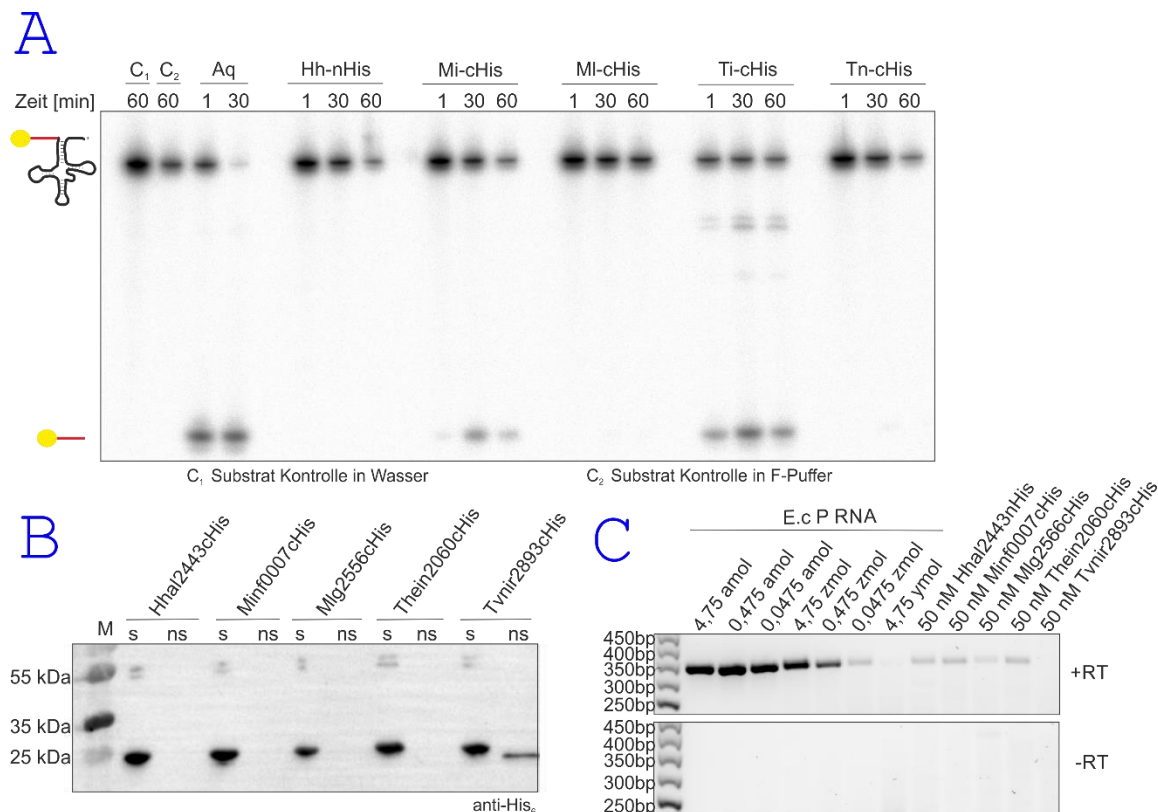


**Abbildung 8: Prozessierung von Akzeptorstamm Varianten.**

Im Anschluss an die Prozessierung wurden die Spaltprodukte auf einem 25%-igen PAA-Gel aufgetrennt. Als Größenstandard wurde ein 16-mer (16), sowie ein Nuklease P1-Verdau des 16-mers (P1) verwendet. Die Spaltung der verschiedenen Substrate wurde mit den Enzymen AaRP, *T. indicus* HARP (Thein), *M. thermautotrophicus* HARP (Mth) und PRORP3 durchgeführt. Unter der Spaltanalyse sind die Substratvarianten schematisch abgebildet, inklusive der Positionen +1/-1, sowie der Flankenlänge (blau/türkis).

Die HARP-Enzyme wurden rekombinant mit einer Histidin-Fusion exprimiert und anschließend chromatographisch gereinigt. Zur Kontrolle wurden die HARP-Präparationen einer Reverse Transkriptase (RT)-PCR-Analyse unterzogen, um eine Kontamination mit *E. coli* RNase P RNA (P RNA) auszuschließen. Anhand sukzessiver Verdünnungen einer *in vitro* transkribierten *E. coli* P RNA konnte eine Detektionsgrenze von 0,0475 zmol bestimmt werden. Außer bei *T. nitratireducens* zeigte sich in allen Präparationen eine geringfügige Verunreinigung mit *E. coli* P RNA in einem Konzentrationsbereich von 47,5 - 475 ymol P RNA (Abbildung 9C). Um zu überprüfen, ob die Kontamination die Spaltanalyse beeinflusst, wurde entsprechend verdünntes *E. coli* Holoenzym (P RNA : P Protein 1:1) in einer Prozessierungsreaktion getestet. Dabei zeigte sich jedoch, dass bereits bei einer Konzentration von 4,75 zmol Holoenzym keine Spaltung mehr detektiert werden kann (vgl. Seite 146). Bei den *in vitro*-Spaltanalysen mit den rekombinant aufgereinigten Enzymen zeigte sich, dass

sowohl das HARP von *M. infernorum* (Mi-cHis) als auch jenes von *T. indicus* (Ti-cHis) in der Lage sind, eine Prä-tRNA effizient an der spezifischen Spaltstelle zu prozessieren (Abbildung 9A). Die HARP-Enzyme aus *H. halophila*, *M. infernorum* und *T. nitratireducens* zeigten in diesem Assay keine Aktivität.



**Abbildung 9: Analyse rekombinanter bakterieller HARPs auf RNase P Funktion**

Die HARP-Proteine aus folgenden Bakterienstämmen wurden für Funktionsstudien rekombinant (mit N- oder C-terminalem Histidin-Fusion) exprimiert und aufgereinigt: *H. halophila* (Hh-nHis, Hha2443cHis), *A. ehrlichii* (MI-cHis, Mlg2556cHis), *T. nitratireducens* (Tn-cHis, Tvnir2893cHis); *M. infernorum* (Mi-cHis, Minf0007cHis) und *T. indicus* (Ti-cHis, Thein2060cHis). (A) Prozessierung der Prä-tRNA<sup>Gly</sup> durch die bakteriellen HARPs. Es ist zu erkennen, dass sowohl *A. aeolicus* (Aq), als auch *M. infernorum* und *T. indicus* Lage sind, die 5'-Flanke endonukleolytisch abzuspalten; Reaktionsbedingungen: 4.5 mM MgCl<sub>2</sub> in 1x F-Puffer bei 37°C. (B) Western Blot Analyse der bakteriellen HARPs die in dem *E. coli* BW-Stamm für die Komplementationsanalyse exprimiert wurden. Alle HARP-Enzyme konnten in der löslichen Fraktion (s) nachgewiesen werden. Allein *T. nitratireducens* zeigte auch ein Signal in der unlöslichen Fraktion (ns). Die Detektion erfolgte mit einem Antiserum gegen den His-Tag. (C) RT-PCR zur Kontrolle möglicher Kontaminationen mit RNase P RNA aus *E. coli* (E.c. P RNA). Als Kontrolle wurde *in vitro* transkribierte E.c. P RNA in einer Verdünnungsserie von 4,75 amol (10<sup>-18</sup>) bis 4,75 ymol (10<sup>-24</sup>) auf das Gel aufgetragen. Für die RT-PCR-Reaktion wurden gereinigte Proteinfractionen verwendet und auf eine Konzentration von 50 nM in dem Reaktionsansatz eingestellt. Es zeigte sich eine Kontamination in allen Proteinfractionen mit Ausnahme von *T. nitratireducens*. Die entsprechende Bande wurde nicht in der Kontrolle ohne Reverse Transkriptase (-RT) detektiert, sodass eine DNA-Kontamination ausgeschlossen werden kann.

Zusätzlich zu den *in vitro*-Prozessierungsanalysen wurden die HARP-Enzyme auch auf ihre Funktion *in vivo* getestet. Hierzu wurde der *E. coli* Stamm BW verwendet, bei dem die Expression der P RNA-Untereinheit strikt Arabinose-abhängig ist [113]. Ein Entzug dieser Kohlenstoffquelle (als Referenz-Kohlenstoff wurde Glucose verwendet) ist durch Depletierung der endogenen RNase P-Aktivität letal für die Zellen. Zur Überprüfung der *in vivo*-Funktionalität wurde der *E. coli* BW Stamm mit Expressionsvektoren für die HARP-Enzyme transformiert und das Zellwachstum auf Medium mit bzw. ohne Arabinose untersucht. Dabei zeigte sich, dass unter permissiven Bedingungen (Glukose-Medium) neben den Positivkontrollen (*E. coli* P RNA und *A. aeolicus* AaRP) ausschließlich das HARP von *T. indicus* in der Lage war, das fehlende Enzym in *E. coli* BW zu komplementieren (Abbildung 10A), obwohl die HARP-Enzyme von *H. halophila*, *M. infernorum*, *A. ehrlichii* und *T. nitratireducens* alle in der löslichen Zellfraktion von *E. coli* BW nachgewiesen werden konnten (Abbildung 9B).

Die Tatsache, dass die untersuchten Bakterienstämme sowohl die Gene für die klassische RNase P als auch ein HARP kodieren, warf die Frage auf, ob die Funktion der klassischen RNase P eventuell in einem dieser Stämme von HARP (zumindest teilweise) übernommen wurde und deshalb die RNA-basierte RNase P bereits an Funktionalität verloren hat. Es wurde daher ebenfalls die Funktion der klassischen RNA-basierten RNase P-Enzyme in den untersuchten Bakterienstämmen adressiert. Dies sollte insbesondere durch *in vivo*-Komplementationstudien geklärt werden. Dafür wurde unter anderem der *Bacillus subtilis* Stamm d7 verwendet, in dem das für die Proteinuntereinheit der RNase P (P Protein) kodierende *rnpA*-Gen unter der Kontrolle eines regulierbaren Xylose-Promoters steht [114]. In Abwesenheit von Xylose kommt es zu einer für den d7-Stamm letalen Inhibition der *rnpA*-Genexpression. Zunächst wurden die *rnpA*-Gene aus der chromosomalen DNA von *A. ehrlichii*, *H. halophila*, *M. infernorum*, *T. indicus* und *T. nitratireducens* mittels PCR amplifiziert und in einen Expressionsvektor für *B. subtilis* inseriert. Dabei stellte sich heraus, dass das P Protein von *M. infernorum* in der NCBI-Datenbank falsch (verlängert) annotiert wurde, sodass in der vorliegenden Arbeit eine verkürzte Variante dieses P Proteins verwendet wurde (vgl. Anhang). Nach Transformation der Expressionsvektoren in *B. subtilis* d7 zeigte sich, dass alle heterologen P

Proteine in der Lage sind, das endogene P Protein von *B. subtilis* unter nicht-induzierenden Bedingungen (Xylose-freies Medium) zu komplementieren (Abbildung 10B). In weiterführenden Studien konnte ferner gezeigt werden, dass die jeweiligen P RNA-Untereinheiten ebenfalls *in vitro* und *in vivo* aktiv sind (*T. indicus* (Fig. S21, Seite 95); bzw. Daten nicht gezeigt).

Dass einige HARP-Enzyme *in vitro* und *in vivo* keine RNase P-Aktivität zeigten, kann mehrere Ursachen haben. Alle getesteten Enzyme stammen aus extremophilen Bakterien. Dadurch ist es möglich, dass die getesteten Reaktionsbedingungen bzw. die zellulären Bedingungen in *E. coli* nicht optimal für die Funktionalität dieser Proteine waren. Es konnte zwar gezeigt werden, dass sich alle HARP-Proteine vorwiegend in der löslichen Zellfraktion von *E. coli* BW befinden (Abbildung 9B), allerdings könnten zusätzliche, in diesem Wirt nicht vorhandene Faktoren für eine Funktion der Enzyme notwendig sein. Auch bei den *in vitro* Analysen könnten für die Enzymaktivität wichtige Bedingungen (Ionen, pH, Reaktionstemperatur usw.) nicht erfüllt gewesen sein. So konnte für *Trypanosoma brucei* gezeigt werden, dass zwei entfernt verwandte Enzyme, mit der gleichen Funktion, nur unter vollkommen unterschiedlichen Reaktionsbedingungen aktiv sind [138]. Weiterhin ergab ein Sequenzvergleich, dass die funktionell äquivalenten Aspartatreste des aktiven Zentrums (D7<sup>Aq</sup>, D138<sup>Aq</sup>, D142<sup>Aq</sup> und D160<sup>Aq</sup>) in den untersuchten HARP-Proteinen konserviert sind (vgl. Seite 98), sodass die Inaktivität kaum auf Mutationen katalytisch relevanter Reste zurückzuführen ist. Allein bei *T. indicus* ließ sich sowohl bei dem Protein-basierten als auch dem RNA-basierten Enzym *in vitro* und *in vivo* RNase P-Aktivität nachweisen, sodass dieses Bakterium tatsächlich zwei strukturell unterschiedliche RNase P-Enzyme besitzt, was ein Novum darstellt.































In der Evolution sind Gen-Duplikationen ein bekannter Mechanismus, um neue Funktionen in einer Zelle zu etablieren [139, 140]. Da der evolutive Selektionsdruck nicht mehr so stark auf ein dupliziertes Gen wirkt, kann es hier häufiger zu Mutationen kommen [141]. Neue Eigenschaften können auch durch horizontalen Gentransfer erworben werden, indem artfremde Gene vom Organismus aufgenommen und ins Genom integriert werden [142, 143]. Da etwa 10% des Genoms von *A. aeolicus* archaealen Ursprungs ist, ist es wahrscheinlich, dass *A. aeolicus* seine RNase P durch horizontalen Gentransfer erhalten hat [144]. Da nahezu alle Organismen, in denen ein HARP-Enzym

identifiziert wurde, Lebensräume mit extremen Bedingungen bewohnen, ist es möglich, dass auch diese Bakterien oder Archaeen das Enzym mittels horizontalem Gentransfer erworben haben. Welche physiologische(n) Funktion(en) die HARP-Enzyme in den jeweiligen Zellen besitzen, ist zurzeit noch unklar. Es wäre auch denkbar, dass die HARP-Gene nach hypothetischer Gen-Duplikation oder horizontalem Gentransfer durch Neofunktionalisierung geänderte Funktionen in der Zelle erlangten [145]. Im Zusammenhang mit der vorliegenden Arbeit konnte in Kooperationen mit der AG Marchfelder (Ulm) und AG Schmitz-Streit (Kiel) gezeigt werden, dass in den Archaeen *H. volcanii* und *M. mazei* die Gene für HARP ohne erkennbaren Phänotyp deletierbar sind (Seite 128-129). Folglich handelt es sich bei HARP in diesen beiden Organismen um ein nicht-essenzielles Enzym. Im Gegensatz dazu war eine Deletion der RNA-basierten RNase P letal, was darauf hindeutet, dass die HARP-Enzyme zumindest nicht das Haupt-RNase P-Enzym in diesen Archaeen darstellen [146]. Es bleibt daher unklar, welche physiologische Funktion das HARP hier ausübt. Es wäre nun interessant herauszufinden, welchen Effekt eine entsprechende Deletion von HARP bzw. der RNA-basierten RNase P in den hier analysierten Bakterien besitzt.

#### **4.5 Strukturaufklärung protein-basierter RNase P**

Die Struktur von Protein-basierten RNase Ps aus *A. thaliana* ist aufgeklärt (Abbildung 3) [50, 110]. Diese Enzyme bestehen aus drei unterschiedlichen Domänen: (1) der PPR-Domäne, welche eine Rolle für die Substraterkennung und -bindung eine Rolle spielt, (2) die NYN-Domäne, welche das aktive Zentrum enthält und (3) der zentralen Domäne, welche ein Zink-Finger Motiv enthält sowie die zwei anderen Domänen miteinander verknüpft [50, 51, 82, 110, 111, 147].

Bioinformatische Sequenzanalysen haben vorhergesagt, dass die prokaryotischen AaRP/HARPs strukturell lediglich aus einer PIN-like Domäne bestehen. Diese Struktur ist ebenfalls charakteristisch für die NYN-Domäne der PRORP-Enzyme [148, 149]. In dieser Arbeit wurden sowohl Analysen zum Oligomerisierungsgrad der AaRP/HARP-Proteine als auch zur Strukturaufklärung durchgeführt. Dazu wurden molekularbiologische, biophysikalische und strukturbiochemische Methoden angewandt.

A			B		
HARP	+ara	+glu	P Protein	+xyl	+glu
	1d	1d		1d	1d
-			-		
Eco			Bsu		
Hha			Hha		
Min			Min		
Mlg			Mlg		
The			The		
Tvn			Tvn		
Aae					

**Abbildung 10: Komplementationsanalysen der bakteriellen HARPs und P Proteine**

Die Komplementationsanalysen wurden in Bakterienstämmen mit reprimierbarer endogener RNase P-Aktivität durchgeführt. Die Expression der endogenen RNase P ist in beiden Stämmen von der Kohlenstoffquelle im Wachstumsmedium abhängig. Zunächst wurden die Bakterienstämme mit den Expressionsvektoren der zu analysierenden heterologen RNase P-Gene transformiert und anschließend auf Agarplatten unter für die endogene RNase P permissiven und nicht-permissiven Bedingungen wachsen gelassen. Als Referenzkohlenstoffquelle für die nicht-permissiven Bedingungen wurde Glucose in der jeweils entsprechenden Konzentration verwendet. *H. halophila* (Hha), *A. ehrlichii* (Mlg), *T. nitratireducens* (Tvn); *M. inferorum* (Min) und *T. indicus* (The). (A) Funktion der bakteriellen HARP-Enzyme *in vivo*. Die Komplementationsanalyse wurde in dem *E. coli* Stamm BW mit Arabinose-abhängiger P RNA-Expression durchgeführt. Der BW-Stamm wurde mit Expressionsvektoren für die heterologen HARP-Enzyme transformiert. Als Negativ-Kontrolle (-) wurde der Leervektor pDG148(+) verwendet, sowie als Positiv-Kontrollen *A. aeolicus* (Aae) RNase P (AaRP) sowie und die homologe *E. coli* P RNA (Eco) exprimiert. Das Wachstum wurde unter induzierenden Bedingungen (+ara) und nicht-induzierenden Bedingungen (+glu) auf LB-Agarplatten nach 1 Tag bei 37°C dokumentiert (B) Funktion der bakteriellen P Proteine *in vivo*. Die Komplementationsanalyse wurde in *B. subtilis* Stamm d7 mit Xylose-abhängiger P Protein-Expression durchgeführt. Der d7-Stamm wurde mit Expressionsvektoren für die heterologen P Proteine-Enzyme transformiert. Als Negativ-Kontrolle wurde der Leervektor pDG148(+) (-) verwendet, als Positiv-Kontrolle die P Proteinuntereinheit von *B. subtilis* (Bsu). Das Zellwachstum wurde unter induzierenden Bedingungen (+xyl) und nicht-induzierenden Bedingungen (+glu) auf LB-Agarplatten nach 1 Tag bei 37°C dokumentiert.

#### 4.5.1 Bestimmung des Oligomerisierungsgrads

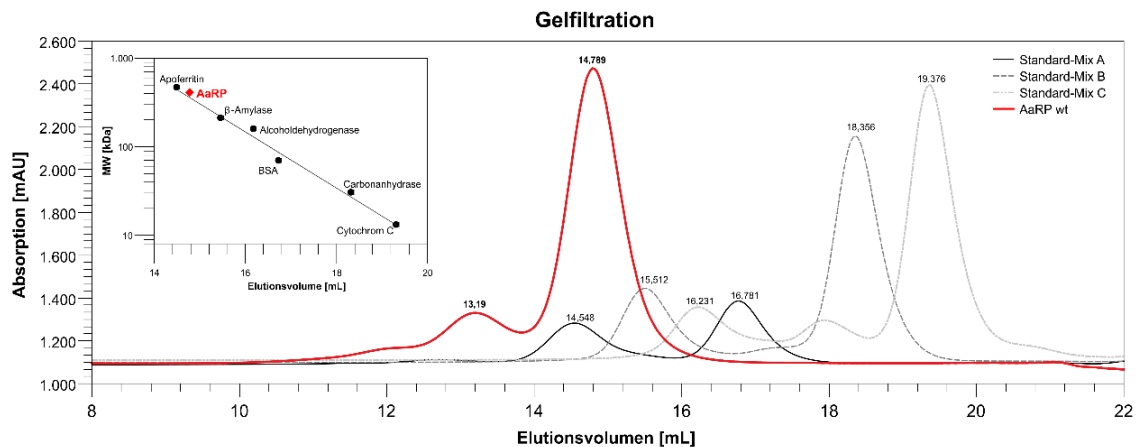
Die Gelfiltration ist eine chromatographische Methode anhand einer Gelmatrix, meistens Sephadex oder Superose, diverse Proteine nach Größen aufzutrennen [150]. In der analytischen Form werden Standardproteine mit bekanntem

Molekulargewicht eingesetzt, um eine Standardgerade zu erhalten welche dann zur Bestimmung von unbekannten Proteingrößen genutzt werden kann.

Zunächst wurde das Elutionsvolumen folgender Kalibrierungsproteine bestimmt: Apoferritin (443 kDa),  $\beta$ -Amylase (200 kDa), Alkoholdehydrogenase (150 kDa), BSA (66 kDa), Carbon Anhydrase (29 kDa) und Cytochrom C (12,4 kDa). Anschließend wurde das Elutionsvolumen gegen die Molekülmasse der Kalibrierungsproteine in logarithmischer Skalierung aufgetragen und die Formel für die Ausgleichsgerade wurde bestimmt. Mittels dem experimentell bestimmten Elutionsvolumen der AaRP/HARP-Proteine und der Geradengleichung für die Gelfiltrationssäule konnte dann die Molekülgröße bzw. der Oligomerisierungsgrad errechnet werden.

Für die RNase P von *A. aeolicus* (AaRP) wurde bereits zuvor gezeigt, dass dieses Protein in einem größeren Komplex vorliegt. So eluierte das native AaRP-Enzym aus *A. aeolicus*-Zellen, das monomer eine Molekülmasse von ~23 kDa besitzt, bei der präparativen Gelfiltration in Fraktionen, die einer Molekülmasse von ungefähr 400 kDa entsprechen (vgl. Fig. S15 (Seite 91)). Um nun herauszufinden, ob mit dem rekombinanten Protein ein vergleichbares Ergebnis erzielt wird, wurde eine analytische Gelfiltration durchgeführt (Superose 6, 10/300 GL, 5 kDa-5 000 kDa). Dabei zeigte sich, dass auch die rekombinante Form bereits früh von der Säule eluiert, vergleichbar mit dem Kalibrationsprotein Apoferritin (443 kDa) (Abbildung 11). Eine kleinere Fraktion von rekombinantem AaRP eluierte bereits noch vor der der Hauptfraktion, was auf eine nochmals höhere Molekülmasse dieser Komplexe hindeutet. Hierbei könnte es sich um unspezifische Aggregate handeln. Das Chromatogramm-Profil mehrerer voneinander unabhängigen AaRP-Präparationen war nahezu identisch. Der Hauptfraktion nach zu urteilen liegt das AaRP-Enzym demnach als homo-oligomerer Komplex mit einer Größe von 400-420 kDa vor. Inwieweit dieser Komplex aus kleineren Einheiten, wie z.B. Dimeren, Trimeren oder Hexameren aufgebaut ist, kann aus den Daten nicht geschlossen werden.





**Abbildung 11: Gelfiltrations-Chromatogramm**

Abgebildet ist ein Lauf der Superose 6 Chromatographie mit den entsprechenden Standard-Proteinen (Grau-Töne), sowie von AaRP (rot). In der oberen linken Ecke befindet sich die korrespondierende Standard-Gerade, welche durch die Elutionsvolumen bestimmt wurde, welche über dem Peak-Maximum angegeben sind.

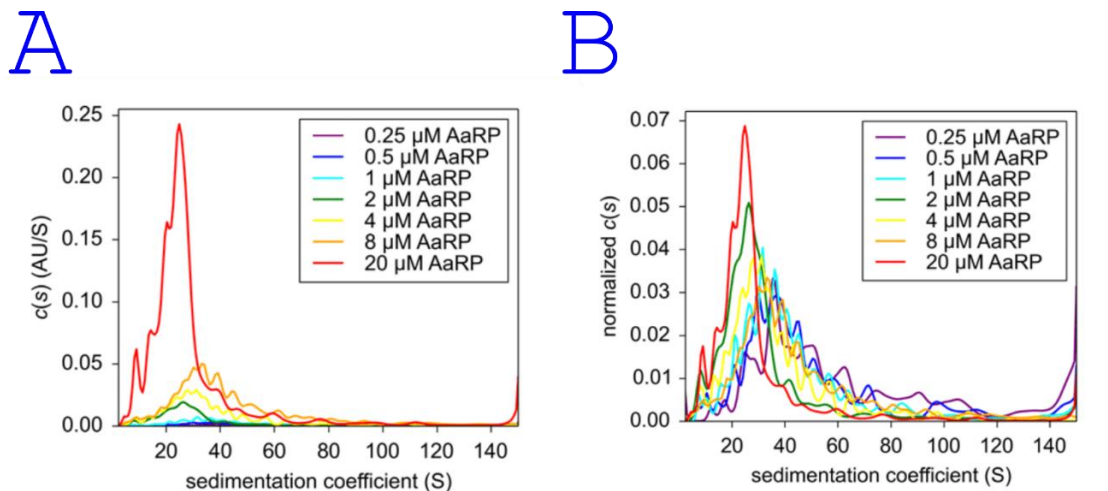
Neben AaRP wurden ein bakterielles und ein archaeales HARP-Proteine mittels Gelfiltration untersucht. Das Elutionsvolumen des bakteriellen HARPs von *T. indicus* lag bei 17,98 mL, was einer Molekülgröße von ~71 kDa entsprach. Für das archaeale HARP von *H. volcanii* konnte ein Elutionsvolumen von 16,85 mL (Molekülmasse ~91 kDa) beobachtet werden. Demnach liegen alle untersuchten AaRP/HARP-Enzyme als Homo-Oligomere vor, jedoch bilden die HARP-Enzyme gegenüber AaRP bedeutend kleinere Komplexe. Eine Strategie von thermophilen Bakterien zur Erhöhung der Proteinstabilität besteht in der Bildung von Oligomeren [151, 152]. Die optimalen Wachstumstemperaturen von *H. volcanii* (42°C) und *T. indicus* (70°C) liegen deutlich unter denen von *A. aeolicus*, welches noch bei Temperaturen bis zu 95°C wachsen kann. Eventuell wird durch die Bildung der höhermolekularen AaRP-Komplexe in *A. aeolicus* eine höhere Thermostabilität erzielt.

#### 4.5.2 Oligomerisierungsgrad Bestimmung durch analytische Ultrazentrifugation

Die analytische Ultrazentrifugation ist eine Methode, die Molekülmasse von Proteinen und Proteinkomplexen mittels eines Dichte-Gradienten zu bestimmen

[153]. Die Analyse erfolgte in Zusammenarbeit mit Dr. Ute Curth (Hannover), welche die Messungen durchführte.

Zunächst wurde AaRP in drei aufeinanderfolgenden Schritten mittels Affinitätschromatographie, Anionenaustauschchromatographie und Gelfiltration aufgereinigt. Nach einer 2,5-stündigen Präinkubation zur Gleichgewichtseinstellung wurden die Proben in einem Dichtegradienten ultrazentrifugiert. Hierbei zeigte sich, dass die Homogenität der AaRP-Präparationen trotz der oben genannten Reinigungsschritte (inklusive Gelfiltration) sehr niedrig war (Abbildung 12). Außerdem konnte beobachtet werden, dass die Komplexe bei höheren AaRP-Konzentrationen kleiner wurden (Abbildung 12B). Möglicherweise handelt es sich bei der Bildung der Oligomere um ein dynamisches System, wobei sich nur kurzzeitig stabile funktionale Komplexe bilden, die meist in Verbindung mit Co-Faktoren vorliegen [154-156]. Eine andere Erklärung ist, dass die Gleichgewichtseinstellung bedeutend länger dauert.



**Abbildung 12: Sedimentationsverhalten von AaRP**

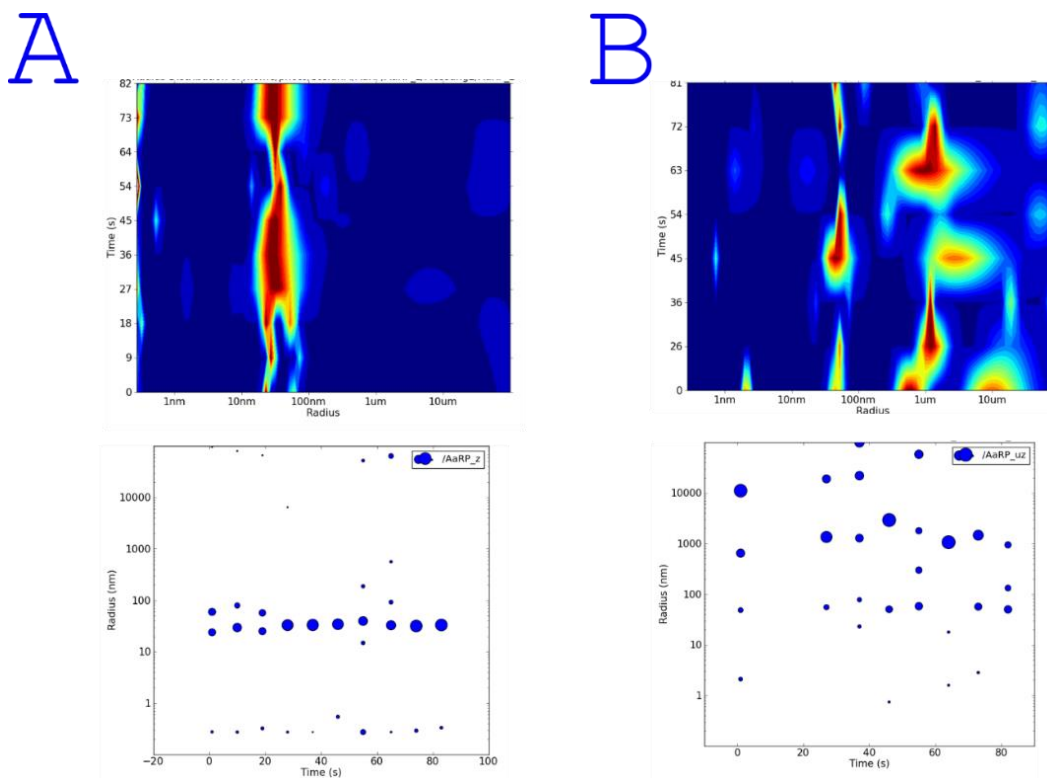
Analytische Ultrazentrifugation von AaRP (Abbildungen von Dr. Curth, Hannover). (A) Sedimentationskoeffizientenverteilungen von AaRP in Abhängigkeit von der Konzentration. (B) Sedimentationskoeffizientenverteilungen nach Normierung der Flächenintegrale.

#### 4.5.3 Größenbestimmung mittels dynamischer Lichtstreuung

Die Dynamische Lichtstreuung (DLS) ist eine Methode zur Bestimmung der Partikelgröße (z.B. von Proteinen und Proteinkomplexen) in einem flüssigen

System mit Hilfe eines Lasers [157]. Die Messungen wurden in Kooperation mit Dr. Stefan Merkl (CrystalsFirst, Marburg) durchgeführt.

Zunächst wurde das rekombinant-exprimierte Protein über Affinitätschromatographie gefolgt von einer Anionentauscherchromatographie gereinigt. Die anschließende Analyse der Proben mittels DLS (CrystalsFirst, Marburg) zeigte, dass die homo-oligomeren Komplexe von AaRP sehr heterogen waren. Die Homogenität der Probe konnte allerdings durch einen zusätzlichen Zentrifugationsschritt verbessert werden, was die Streuung der Partikel in der DLS-Messung verringerte. Die dadurch bestimmten Partikelgrößen hatten einen Radius von 10-100 nm, vereinzelt konnten jedoch auch deutlich kleinere Partikel (~1 nm Radius) detektiert werden.



**Abbildung 13: Größenanalyse von AaRP mittels Dynamischer Lichtstreuung**

Darstellung der Ergebnisse der Dynamischen Lichtstreuung in Form einer *Heat-Map* (obere Abbildungen). Im unteren Bereich sind die entsprechend ermittelten Partikel-Radien abgebildet (Dr. Merkl, Marburg). (A) Chromatographisch aufgereinigte AaRP-Präparation mit anschließender Zentrifugation. (B) Chromatographisch aufgereinigte AaRP-Präparation ohne Zentrifugation

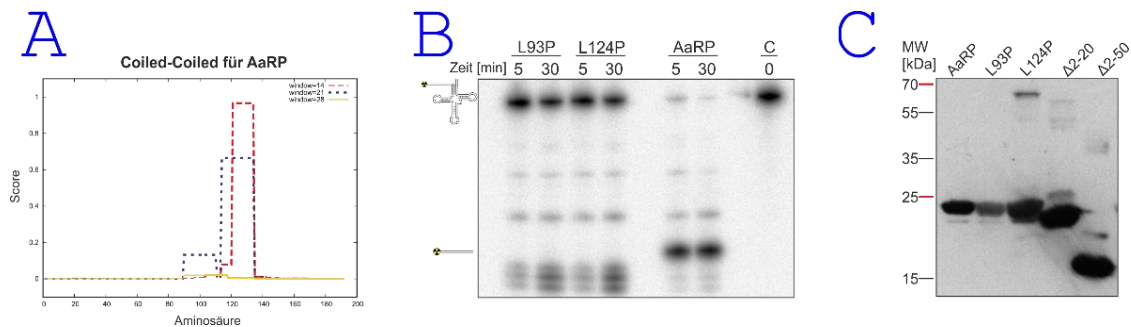
#### 4.5.4 Strukturanalyse durch Aminosäureaustäusche

Die vorangegangenen Ergebnisse zeigen, dass AaRP vorwiegend als größerer homo-oligomerer Komplex vorliegt. Durch den gezielten Austausch einzelner Aminosäuren wurde versucht, die für die Komplexbildung verantwortlichen Strukturelemente von AaRP näher zu charakterisieren.

In ersten bioinformatischen Analysen wurden für zwei Sequenzbereiche von AaRP Proteinstrukturmotive (*coiled-coiled* Domänen) vorhergesagt, welche beispielsweise in Transkriptionsfaktoren Protein-Protein Interaktionen vermitteln (Abbildung 14A). Im Folgenden wurden zwei Leucine (L93 und L124) innerhalb von  $\alpha$ -Helices der vorhergesagten *Coiled-Coil*-Motive durch Prolin ersetzt, um eine mögliche Protein-Protein-Interaktion zu inhibieren. Zunächst wurde die katalytische Aktivität der AaRP-Varianten in einem Prä-tRNA-Spaltungsassay bestimmt (Abbildung 14B). Sowohl die L93P- als auch die L124P-Mutation inhibierte die Prä-tRNA-Prozessierungsaktivität der AaRP-Proteine vollständig (die unspezifische Spaltung im Hintergrund ist auf das im Puffer vorhandene BSA zurückzuführen). Zur strukturellen Analyse der AaRP-Varianten wurde ein Western Blot und eine Gelfiltration durchgeführt. In der Western Blot-Analyse konnte ausschließlich bei AaRP (L124P) eine diskrete höhermolekulare Bande immundetektieren werden (Abbildung 14C). Bei der Gelfiltrationsanalyse konnte für beide AaRP-Varianten kein definiertes Elutionsmaximum bestimmt werden, was auf eine heterogene Größenverteilung der Proteine hindeutet, oder aus ein denaturiertes bzw. nicht-gefaltetes Protein. Es ist demnach möglich, dass es sich bei der höhermolekularen Bande im Western Blot um unspezifische Aggregatbildung handelt und sowohl AaRP (L124P) als auch AaRP (L93P) denaturiert ist. Wenn durch die jeweilige Mutation die Oligomerisierung inhibiert worden wäre, so hätte die monomere Form von AaRP bei der Gelfiltration bei ~23 kDa in Form eines definierten Peaks eluieren müssen. In weiteren Experimenten wurden unterschiedlich lange Bereiche des N-terminalen Endes von AaRP deletiert (19 bzw. 49 Aminosäuren). In diesem Sequenzabschnitt befindet sich zwar der katalytisch relevante Aspartatrest an Position 7, jedoch sollte mit diesen Analysen die RNA-Bindung bzw. die Oligomerisierung von AaRP charakterisiert werden. In der Western Blot-Analyse konnten sowohl bei AaRP ( $\Delta$ 2-20) als auch bei AaRP ( $\Delta$ 2-50) schwache höhermolekulare

Oligomere/Aggregate immunodetektiert werden (Abbildung 14C), so dass auf eine Gelfiltration verzichtet wurde.

Anhand dieser Ergebnisse spielt der N-Terminus von AaRP, ebenso wie die vorhergesagten *Coiled-coiled*-Domänen, scheinbar keine Rolle bei der Oligomerisierung.



**Abbildung 14: Analyse der *Coiled-Coil* Vorhersage für AaRP**

(A) Bioinformatische Analysen sagen für zwei Bereiche von AaRP *coiled-coiled* Motive voraus. Der Score gibt die Wahrscheinlichkeit für die Vorhersage an. (B) Aktivitätsbestimmung von AaRP (L93P) und AaRP (L124P). AaRP (Wildtyp) diene als Positivkontrolle. Als Substrat wurde 5'-markierte Prä-tRNA<sup>Gly</sup> aus *T. thermophilus* verwendet. In der Kontrolle (C) wurde das Substrat in Reaktionspuffer für 30 min inkubiert. (C) Western Blot-Analyse der verschiedenen AaRP-Varianten AaRP (Wildtyp), AaRP (L93P), AaRP (L124P), AaRP ( $\Delta$ 2-20) und AaRP ( $\Delta$ 2-50). Die rekombinant aufgereinigten Proteine wurden auf einem 15% SDS-Gel aufgetrennt und nach dem Transfer auf eine PVDF-Membran mittels AaRP-spezifischen Antikörpern immunodetektiert.

#### 4.5.5 Strukturanalyse mittels Circular dichroismus-Spektroskopie

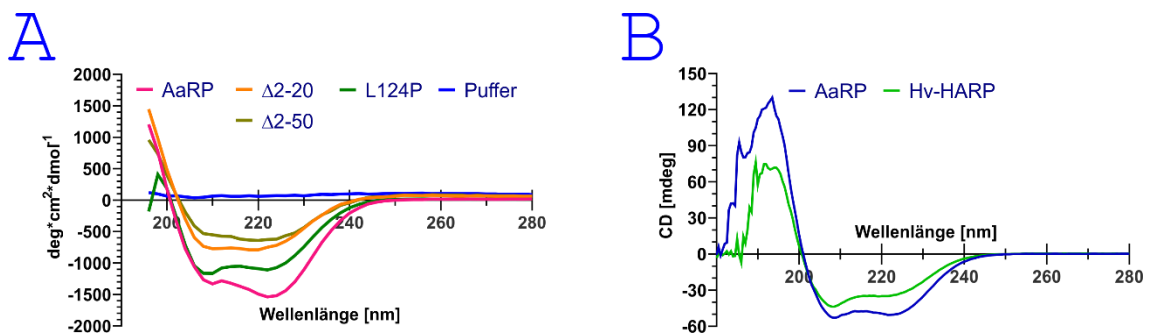
Die Circular dichroismus-Spektroskopie (CD-Spektroskopie) erlaubt die Strukturaufklärung chiraler Moleküle unter Verwendung von zirkular polarisiertem Licht. Diese Methode wird oft zur Analyse von Sekundärstrukturen in Proteinen verwendet. Dabei kann anhand des erhaltenen Spektrums eine Aussage über die Anteile an  $\alpha$ -Helices und  $\beta$ -Faltblättern bestimmt werden. Sollten die Proteine unstrukturiert oder denaturiert vorliegen, kann dies anhand der Spektren erkannt werden [158]. Die Proteine wurden Affinitätschromatisch gereinigt.

Die einzelnen AaRP/HARP-Varianten wurden optisch vermessen und mit dem Spektrum von AaRP (Wildtyp) verglichen. Die CD-Analyse zeigte, dass AaRP in einer fast ausschließlich  $\alpha$ -helikalen Form vorliegt, was an den zwei Absorptionsminima bei  $\sim 210$  nm und  $\sim 225$  nm zu erkennen ist (Abbildung 15). Die AaRP-Varianten mit deletiertem N-terminalen Ende und AaRP (L124P)

weisen ebenfalls einen hohen  $\alpha$ -helikalen Anteil auf (Abbildung 15A). Demzufolge ist davon auszugehen, dass die Mutationen/Deletionen keine größeren strukturellen Veränderungen hervorgerufen haben.

Überraschend war das Ergebnis der CD-Analyse von dem archaealen *H. volcanii* HARP-Enzym. Den Daten zufolge weist dieses Protein gegenüber AaRP ein deutlich höherer Anteil an  $\beta$ -Faltblattstrukturen auf, was an dem fehlenden Absorptionsminimum bei 225 nm zu erkennen ist. Das heißt das Protein enthält sowohl  $\alpha$ -Helices als auch  $\beta$ -Faltblattstrukturen (Abbildung 15B).

Die Ergebnisse der CD-Analyse von AaRP (L124P) deuten im Gegensatz zu den Daten der Gelfiltrationsanalyse darauf hin, dass die Helix, in welcher Leucin 124 lokalisiert ist, für die Gesamtstruktur von AaRP nicht essenziell ist. Auch konnte durch die CD-Analysen gezeigt werden, dass die Gesamtstruktur von AaRP und archaealem HARP verschieden ist. Dies könnte einen Einfluss auf die katalytische Aktivität und Substratbindung haben, so dass sich die physiologischen Funktionen von AaRP und HARP unterscheiden könnten.



**Abbildung 15: CD-Analyse von AaRP und Hv-HARP**

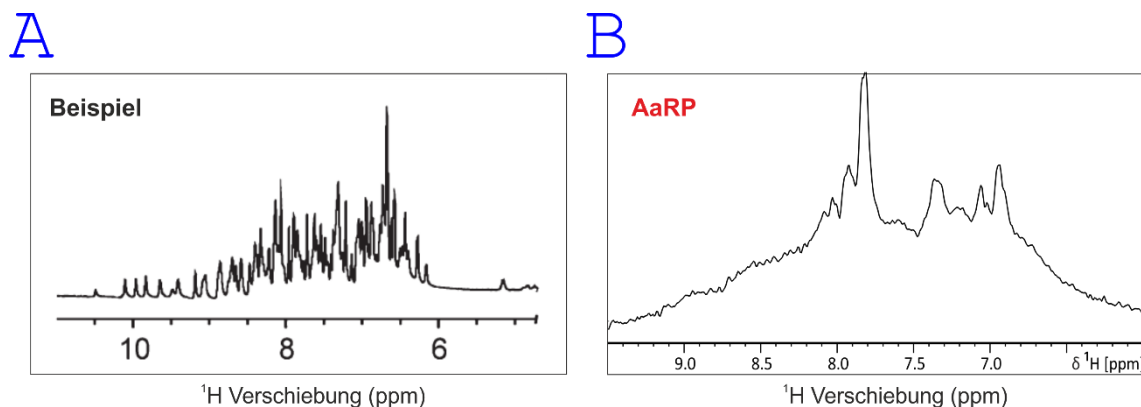
Abgebildet sind die CD-Spektren von AaRP-Varianten und archaealem HARP aus *H. volcanii*. Die gemessene Ellipzität der Proteine ist gegen die Wellenlänge aufgetragen. (A) CD-Spektren von AaRP (Wildtyp), AaRP (L124P), AaRP ( $\Delta$ 2-20) und AaRP ( $\Delta$ 2-50). (B) CD-Spektren von AaRP (Wildtyp) und *H. volcanii* HARP.

#### 4.5.6 Strukturaufklärung mittels Kernspinresonanzspektroskopie

Die Kernspinresonanzspektroskopie (NMR) ist eine Methode, um unter anderem die Struktur von Makromolekülen aufzuklären. Dabei werden graphisch alle Protonen und deren chemische Verschiebung (ppm) aufgezeigt. Bei Proteinstrukturen sind dies die Amidprotonen aus den Peptidbindungen, welche eine Verschiebung von 6 - 10 ppm besitzen [159].

Das rekombinante AaRP wurde zunächst affinitätschromatisch gereinigt, gefolgt von einem Ionenaustauscher. Anschließend wurde das Protein an die Goethe-Universität Frankfurt (AG Wöhnert) gesendet. Das 1D  $^1\text{H}$  Spektrum wurde an einem 600 MHz Spektrometer aufgenommen bei 316.5 K. Als Pulssequenz wurde ein *hs11echo* verwendet mit 1024 Scans. Die Proteinkonzentration betrug 186  $\mu\text{M}$  (4,28 mg/mL).

Bei der Aufnahme des NMR 1D  $^1\text{H}$  Spektrums ließ die Verschiebung der Atome keine klaren Peaks erkennen (beispielhaft dargestellt in Abbildung 16A). Dieses Ergebnis deutet darauf hin, dass AaRP (~22,5 kDa) als größerer Komplex vorliegt und verifiziert nochmals die Daten aus den vorherigen Strukturanalysen (Gelfiltration, Analytische Ultrazentrifugation, Dynamische Lichtstreuung). Auch lässt sich aus dem Spektrum schließen, dass es sich um ein gefaltetes Protein handelt, da ungefaltete Proteine eine Verschiebung bei 7,5 - 8,5 ppm zeigen.



**Abbildung 16: 1D  $^1\text{H}$  Spektrum von AaRP**

Dargestellt sind die Ergebnisse einer NMR-Spektroskopie-Analyse. (A) Exemplarisch ist die Verschiebung der Protonen für ein Protein (CtBP-THAP, verändert nach Kwan *et al.* 2011 [159]) dargestellt. (B) Spektrum von AaRP (Dr. Carolin Hacker, Goethe-Universität Frankfurt).

#### 4.5.7 Kristallisation der RNase P von *Aquifex aeolicus*

Die Kristallisation des Proteins wurde sowohl durch MarXtal (Uni Marburg), als auch in Kooperation mit der AG Reuter (Uni Marburg) durchgeführt. Bei MarXtal wird zum Ansetzen der Kristallisationsplatten ein automatisiertes System verwendet, in der AG Reuter erfolgte eine manuelle Erstellung der Platten. Das erste Ziel war es eine Bedingung zu finden in der die RNase P von *A. aeolicus*



Kristalle bildet und diese im Anschluss zu optimieren, sodass eine Strukturanalyse der Kristalle mittels Röntgenbeugungsexperimenten möglich ist.

Das Protein wurde rekombinant in *E. coli* Rosetta (DE3) exprimiert, anschließend über Affinitäts-, gefolgt von Ionenaustauschchromatographie gereinigt. Für das Ansetzen der Kristallisationsplatten wurden zunächst kommerziell erhältliche Pufferlösungen verwendet. Dabei wurde im kleinen Maßstab die manuelle Variante gewählt, sowie im großen Maßstab bei MarXtal. Insgesamt wurden >1000 Bedingungen untersucht, um eine mögliche Kristallbildung zu detektieren. Der Erfolg wurde in beiden Fällen photographisch dokumentiert, auch bei MarXtal.

Bei den ersten Tests konnte nach 126 Tagen ein Kristall unter der Bedingung des Screens AmSO<sub>4</sub> (Qiagen) in Position H11 (0,1 M MES pH 6,5; 3,5 M Ammoniumsulfat; 1% (w/v) 2-Methyl-2,4-pentandiol) (Abbildung 17, erste Reihe) und unter der Bedingung des Screens JSCG Core I (Qiagen) in Position F11 (0,1 M Natriumacetat pH 4,5; 0,2 M Lithiumsulfat; 50% (v/v) PEG-400) detektiert werden. Bei der Aufnahme eines Diffraktionsmusters für den Kristall AmSO<sub>4</sub>-H11 zeigte sich, dass es sich um ein Kleinmolekül (Salz) handelt (Abbildung 17, oberste Reihe). Die Puffer-Bedingungen von dem Kristall aus Core I-F11 wurden als Ausgangspunkt für eine weitere Optimierung genutzt. Dabei wurden der pH-Wert, die Salzkonzentration und das Präzipitans variiert. Bereits nach 10 Tagen Inkubation bei 18°C konnten auf diese Weise kleine Kristalle identifiziert werden. Allerdings konnte auch bei längerer Inkubation kein weiteres Wachstum der Kristalle beobachtet werden, so dass sie nicht für weitere Experimente verwendet wurden.

Als nächstes wurden Impfkristall-Experimente mit einem sogenannten *Mutter*-Kristall durchgeführt. Dazu wurde ein Kristall (Bedingung Core I F11) mechanisch zerstört und die entstanden Fragmente anschließend mit einem Pferdehaar durch einen Kristallisationstropfen gezogen. Zur Optimierung wurden die Kristallisationsbedingungen (pH, Salzbedingungen) dabei variiert. Überraschenderweise konnten bereits nach wenigen Stunden kleine reisförmige Kristalle beobachtet werden, welche sich in einer Spalte (Spalte 3, Natrium-Citrat Puffer pH 5,6; 50% PEG-400; 0,05 M bis 0,20 M Lithiumsulfat) der 24-Vertiefungen-Platte befanden. Größere Kristalle wurden bereits nach 33 Tagen detektiert. Für die verschiedenen Kristalle aus dem Impfkristall-Ansatz wurde

ebenfalls ein Diffraktionsmuster aufgenommen (Abbildung 17, Core-F11 B3, C2, C5). Dabei konnte eine Mischung aus Kleinmolekül und Protein identifiziert werden (B3 und C5) sowie ein nicht genau definierbares Muster für C2. Da eine Mixtur aus Kleinmolekül und Protein nicht erfolgsversprechend war, wurden diese Kristalle ebenfalls verworfen. Bei den reiskornförmigen Kristallen könnte es sich um spezifische AaRP-Kristalle handeln und weitere Impfkristall-Experimente sind in Planung.

Das Kristallwachstum von AaRP ist sehr langsam. In Abbildung 18 ist auf der linken Seite eine Auswahl verschiedener Kristalle dargestellt, welche bei dem automatisierten Screen identifiziert wurden. Dabei handelt es sich um folgende Bedingungen Core I-A3 (0,05 M Lithiumsulfat; 0,05 M Natriumsulfat; 0,05 M TRIS pH 8,5; 30% PEG-400), Core I-E9 (0,2 M tri-Lithiumsulfat; 20% PEG-335), Core IV-G3 (0,2 M Lithiumsulfat, 0,1 M MES pH 6, 20% 1,4-Butanediol), Core IV-G9 (0,01 M Eisen(III)-chlorid; 0,1 M tri-Natriumcitrat pH 6,5; 10% Jeffamine® M-600), Morpheus II-E11 (90 mM LiNaK; 0,1 M Puffersystem 5 pH 7,5; 30% Precipitant-Mix 7) mit einem Wachstum von maximal 546 Tagen (78 Wochen). Im Puffer von Core IV-A2 (1,2 M Natrium-di-hydrogenphosphat; 0,8 M Kalium-di-hydrogenphosphat; 0,2 M Lithiumsulfat; 0,1 M Glycin) konnten Kristalle nachgewiesen werden, welche der Form des Kristalls im Proteintropfen entsprachen. Dies ist ein Hinweis darauf, dass es sich dabei um Kleinmoleküle handelt und diese Kristalle sich nicht für weitere Untersuchungen eignet. Insgesamt konnten ~20 Kristalle unter den >1000 Bedingungen identifiziert werden, in denen ein Kristallwachstum stattgefunden hat. Lediglich in einer dieser Bedingungen (Core II-H2, 0,2 M Lithiumsulfat; 0,1 M Natriumacetat pH 4,5; 30% PEG-8000; 18°C) konnte bereits ein kleiner Kristall beobachtet werden, welcher auch minimal an Größe zunahm. Allerdings wurde dieser Kristall durch das Präzipitat auseinandergebrochen. Erst nach weiteren 83 Tagen (~12 Wochen) konnten weitere Kristalle identifiziert werden. Eine Testung dieser Kristalle war aufgrund der kleinen Größe nicht möglich.

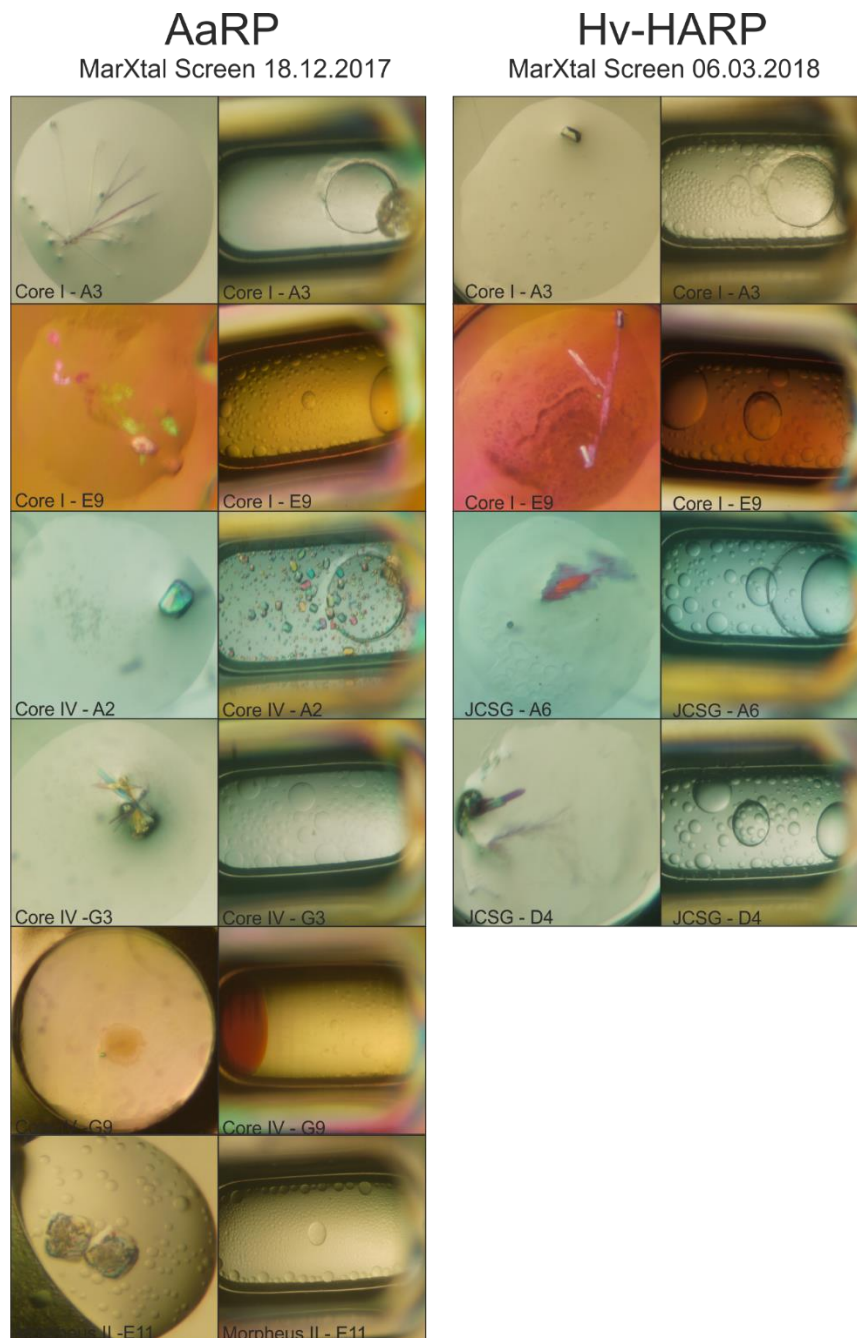
Das langsame Kristallwachstum kann aufgrund der unterschiedlichen Oligomerisierungsgrade stattfinden. Falls es sich um ein dynamisches Gleichgewicht zwischen der monomeren und multimeren Form handelt, worauf die Ergebnisse der analytischen Ultrazentrifugation (4.5.2) hindeuten, könnten variierende Konzentrationsverhältnisse bzw. eine langsame

Gleichgewichtseinstellung den Kristallisationsprozess erschweren. Die kann dazu führen, dass das Kristallwachstum zu früh zum Erliegen kommt und sich Kleinmoleküle durchsetzen.

	Start	Bevor	Kristall	Messung	X-Ray	
AmSO <sub>4</sub> (H11)	 18.12.17	 21.02.18	 23.04.18	 22.08.18	 22.08.18	Salz
Morpheus II (A3)	 18.12.17	 14.06.18	 30.07.18	 22.08.18	 22.08.18	Salz
Morpheus II (A7)	 18.12.17	 30.07.18	nicht vorhanden	 22.08.18	 22.08.18	Protein
Core-F11 (B3)	 29.08.18	nicht vorhanden	 01.10.18	 30.10.18	 31.10.18	Mix
Core-F11 (C2)	 29.08.18	nicht vorhanden	 01.10.18	 30.10.18	 31.10.18	?
Core-F11 (C5)	 29.08.18	nicht vorhanden	 16.10.18	 30.10.18	 31.10.18	Mix
H11-B4 (LiSO <sub>4</sub> )	 08.05.18	 17.05.18	nicht vorhanden	 17.06.19	 01.07.19	?

**Abbildung 17: Kristalle von AaRP mit Diffraktionsanalysen**

In der obigen Abbildung sind ausgewählte Kristalle von AaRP aufgeführt (18°C). Dabei ist der Kristallisationstropfen am Beginn (Start), sowie vor der Detektion des Kristalls (Bevor) abgebildet. Das erste Auftreten des Kristalls (Kristall) wurde ebenfalls festgehalten, zusätzlich wurde der Kristall vor der Aufnahme des Diffraktionsmusters (Messung) nochmals dokumentiert. Das erhaltende Diffraktionsmuster (X-Ray) ist mit dem Aufnahmedatum abgebildet. Die Bedingung der Kristallisation ist der Bilderreihe vorangestellt, sowie eine Interpretation am Ende. Dabei handelt es sich bei Salz um ein Kleinmolekül, bei Mix um ein nicht genau ein zuordnenen Kristall (Kleinmolekül und Protein), sowie um Proteine. Bei einem Fragezeichen (?) ist die Einordnung des Kristalls nicht möglich.



**Abbildung 18: Kristalle von AaRP/HARP des MarXtal-Screens**

Kristalle von AaRP (Spalte links) und *H. volcanii* HARP (Spalte rechts). In jeder Spalte sind die Kristalle (links) und die entsprechenden Pufferreservoirs (rechts) abgebildet. Die Kristallisationsbedingungen sind jeweils in der unteren Hälfte des Bildes angegeben. Es ist zu erkennen, dass in manchen Bedingungen schon das Pufferreservoir Kristalle aufweist (Core IV, A2, AaRP (1,2 M Natriumhydrogenphosphat; 0,8 M Kalium-di-hydrogenphosphat; 0,2 M Lithiumsulfat; 0,1 M Glycin)).

#### 4.5.8 Kristallisation des HARPs von *Haloferax volcanii*

Die Kristallisationsanalyse von HARP aus *H. volcanii* (Hv-HARP) erfolgte wie für AaRP beschrieben (4.5.7). Manuell und automatisiert wurden insgesamt ca. 1000 Pufferbedingungen untersucht. Im Gegensatz zu AaRP bildet Hv-HARP keine großen Komplexe. So entspricht das Elutionsvolumen von Hv-HARP bei der Gelfiltration einer Molekülgröße von ~91 kDa (AaRP eluiert bei ~420 kDa). Anhand der CD-Spektren weist Hv-HARP im Vergleich zu AaRP auch einen höheren Anteil an  $\beta$ -Faltblattstrukturen auf (4.5.5). Folglich könnten sich die Kristallisationseigenschaften der beiden Proteine deutlich voneinander unterscheiden.

In den manuell erstellten Kristallplatten konnten einige Kristalle identifiziert werden. Anhand der Morphologie wurde zunächst von Proteinkristallen ausgegangen, so kam es bei einer Optimierung der jeweiligen Pufferbedingungen auch zu einem Fortschreiten des Wachstums hin zu kubischen, glatten Kristallen. Allerdings ergab eine anschließende Analyse, dass es sich bei diesen Kristallen ausschließlich um Salzkristalle gehandelt hat.

Abbildung 18 zeigt auf der rechten Seite eine Auswahl an Kristallen, welche in unterschiedlichen Bedingungen identifiziert wurden. Zusätzlich wurde das Pufferreservoir mit abgebildet, um auszuschließen, dass es sich um Pufferbedingte-Kleinmoleküle handelt. Die Bedingungen mit Kristallwachstum sind: Core I-A3 (0,05 M Lithiumsulfat; 0,05 M Natriumsulfat; 0,05 M TRIS pH 8,5; 30% PEG-400), Core I-E9 (0,2 M Ammoniumsulfat; 20% PEG-3350), JCSG-A6 (0,2 M Lithiumsulfat; 0,1 M Natriumphosphatcitrat pH 4,2; 20% PEG-1000) und JCSG-D4 (0,2 M Lithiumsulfat; 0,1 M Natriumacetat pH 4,5; 30% PEG-8000). Die Inkubationstemperatur lag bei 18°C. Eine zusätzliche Testung dieser Kristalle konnte im Rahmen dieser Arbeit nicht durchgeführt werden und steht noch aus. Neben den Kristallen, welche für die Beugungsexperimente in Betracht gezogen werden, wurden Sphäroliten identifiziert (Abbildung 19). Diese Art der Kristallbildung konnte in drei unterschiedlichen Bedingungen gefunden werden. Dabei handelt es sich um AmSO<sub>4</sub>-G8 (0,5 M Lithiumchlorid; 1,6 M Ammoniumsulfat), Core II-G12 (0,2 M Ammoniumsulfat; 0,1 M Natriumphosphatcitrat pH 4,2; 20% PEG-300; 10% Glycerol) und Anions-C6 (0,1 M Natriumacetat pH 4,6; 1,2 M di-Natriummalonat). Die ersten positiven

Resultate erfolgten nach 92 Tagen (13 Wochen). Die Methode der Impfkristalle sollte auch bei diesem Protein angewendet werden, um die Zeit der Kristallbildung möglichst zu verkürzen.



**Abbildung 19: Kristalle und Sphäroliten von Hv-HARP**

Auswahl an Sphäroliten und Kristallen von Hv-HARP. Die Bedingung (rechts oben), als auch das Datum der Aufnahme (rechts unten) sind angegeben. Größere Kristalle (Morpheus I, C8) kommen nur einzeln vor, wohingegen Sphäroliten den gesamten Tropfen ausfüllen (AmSO<sub>4</sub>, H8; Core II, G12; Anions, C6). Die Wachstumszeit betrug zwischen 92-171 Tage für die Sphäroliten, sowie 282 Tage für die kleinen stäbchenförmigen Kristall.

#### 4.5.9 Diskussion der Kristallisation

Zurzeit ist es noch nicht möglich Kristallisationsbedingungen vorherzusagen, so dass man zunächst experimentell eine Bedingung finden muss, in derer sich Kristalle bilden, welche dann im nächsten Schritt optimiert werden müssen [160, 161]. Mit Abschluss der vorliegenden Arbeit war es noch nicht möglich, definierte Kristallisationsbedingungen für die beiden Enzyme AaRP aus *A. aeolicus* und HARP aus *H. volcanii* anzugeben. Es konnten bei den oben beschriebenen initialen systemischen Testverfahren zwar Reaktionsbedingungen für eine Kristallisation der Proteine gefunden werden, diese müssen aber in nächsten Schritten noch optimiert werden. Die Durchführung automatisierter Screens bei MarXtal bietet die Möglichkeit, eine Vielzahl unterschiedlicher Bedingungen zu testen. Diese initialen Testverfahren liefern aber häufig nur Kristalle von niedriger Diffraktionsqualität und müssen im weiteren Verlauf noch optimiert werden [162, 163]. Die Proteinkonzentration spielt bei der Optimierung eine maßgebliche Rolle. Eine zu geringe Konzentration führt zu klaren Tropfen und einer Untersättigung, wohingegen es bei einer zu hohen zu einer Übersättigung kommen kann, bei derer das Kristallwachstum in einem Plateau endet. [163]. Bei dem Versuch der Kristallisation von AaRP und Hv-HARP wurden Konzentrationen von 6,6 mg/mL (AaRP) bzw. 9,9 mg/mL (Hv-HARP) gewählt. In

beiden Fällen konnte eine Präzipitation von über 50% der Tropfen beobachtet werden. Da bei beiden Bedingungen noch eine hohe Anzahl an klaren Tropfen beobachtet werden konnten, wäre eine Erhöhung der Proteinkonzentration für einen neuen Ansatz ratsam, um sowohl mehrere Kristalle zu erhalten als auch einen Geschwindigkeitsprofit. Bereits vorhandene Kristalle können ebenfalls für die Optimierung genutzt werden in Form von *Microseeding*. Dies wurde für AaRP bereits durchgeführt, erbrachte aber keinen Erfolg. Die erhaltenen Kristalle wuchsen bereits nach wenigen Tagen, allerdings nahm die Größe mit der Zeit nicht zu. Mit diesen Kristallen könnten erneut verschiedene Bedingungen überprüft werden [164], so wie die Zugabe verschiedener Additive getestet werden. Durch das langsame Wachstum der Kristalle wäre es auch möglich, dass man Seeding durch Öle und Sphäroliten oder Präzipitanten nutzt. So konnte für das Enzym Cellobiose Dehydrogenase (CDH) Kristalle in Öl-Tropfen detektiert werden, welche hoch mit Protein angereichert waren. Allerdings erst nach einem Jahr [165], was einen Vorteil in der Größe und einen Nachteil in der Geschwindigkeit birgt. Eine weitere Möglichkeit, welche aufgrund der vorhandenen Tropfen mit Präzipitat naheliegt, ist das Seeding mit diesen. Bereits die T7 RNA Polymerase konnte auf diese Weise kristallisiert werden [166-168]. Eine Evaluation der Tropfen von AaRP und Hv-HARP im Hinblick auf denaturiertes Protein, sowie amorphe/nicht-amorphe Präzipitation scheint eine gute Möglichkeit um Bedingungen zu finden, bei derer sich das Seeding mit bereits erhaltenen kleinen Kristallen lohnt. Auch die Zugabe von Glycerol, welches ebenfalls bei der Kristallisation der T7 RNA Polymerase eingesetzt wurde, kann die Stabilität der Proteinkristalle erhöhen, sowie als Kryoprotektivum dienen [168]. Bei den meisten Bedingungen, in denen bereits kleine Kristalle erhalten wurden, zeigte sich ein hoher Anteil an Glycerol beziehungsweise PEG-Variationen, welches ähnlich zu Glycerol ebenfalls ein Kryoprotektivum darstellt. Ein weiterer experimenteller Ansatz zur Kristallisation von AaRP/HARP könnte die gezielte Proteinmodifikation sein. So zeigte sich, dass eine aktive Form des MRPP3 (Untereinheit der humanen mitochondrialen RNase P) in dieser Form nicht kristallisierbar war. Erst eine Deletion des N-terminus zeigte positive Resultate [52]. Da MRPP3 wie AaRP/HARP eine Metallonuklease-Domäne enthält, könnte dieser Ansatz ebenfalls erfolversprechend sein.



## 5 Ausblick

In dieser Arbeit konnte die RNase P von *Aquifex aeolicus* (AaRP) identifiziert und charakterisiert werden. Erstaunlicherweise handelte es sich nicht um die klassische bakterielle RNase P, aufgebaut aus einer katalytischen RNA und einer Proteinuntereinheit, sondern um eine rein Protein-basierte Form. Solch eine Form war bis dahin nur aus höheren Organismen bekannt. Die ersten Charakterisierungen dieses neuartigen Enzyms sind bereits durchgeführt worden. Eine der größten Fragen, welche auch durch diese Arbeit noch nicht beantwortet werden konnte, ist die der Substratbindung. Zwar konnte in enzymkinetischen Analysen mit Prä-tRNA-Varianten gezeigt werden, welche Substratdeterminanten für die Substraterkennung wichtig sind, allerdings konnten noch keine aussagekräftigen Bindungsstudien durchgeführt werden. Zwar zeigten initiale Experimente eine Bindung des Proteins mit dem Substrat, jedoch konnten noch keine Parameter bestimmt werden. Neben der biochemischen Charakterisierung sollte auch die Aufklärung der Struktur weitere Erkenntnisse liefern. Zwar konnten im Rahmen dieser Arbeit initiale Bedingungen für eine Kristallisation gefunden werden, die gezüchteten Kristalle waren für weiterführende Analysen jedoch noch zu klein. Auf dieser Grundlage kann jedoch eine weiterführende Optimierung der Kristallisationsbedingungen stattfinden. Eine weitere offene Frage bleibt, in welchem Oligomerisierungszustand das Enzym aktiv vorliegt und welche Proteinstrukturen an der Oligomerisierung beteiligt sind. Erste Analysen zeigten noch keine eindeutige Reduktion der Oligomerisierung des Proteins.

In bioinformatischen Analysen konnten AaRP-homologe Enzyme (HARP) in verschiedenen Prokaryoten identifiziert werden. Im Gegensatz zu *A. aeolicus* verfügen diese Bakterien und Archaeen aber alle noch über eine RNA-basierte Form der RNase P. Erste Analysen zu den archaealen HARP-Enzymen deuten darauf hin, dass sie nicht essenziell und folglich auch nicht die Hauptform der RNase P in diesen Organismen sind. Um Hinweise zur physiologischen Funktion der HARP-Enzyme zu bekommen, sollten weitere Analysen durchgeführt werden. Beispielsweise sollten Transkriptom-Sequenzierungen (RNA-Seq.) dabei helfen, mögliche Substrate von HARP in der Zelle zu identifizieren. Allerdings sind für viele der identifizierten Bakterien/Archaeen noch keine

gentechnischen Methoden etabliert, so dass beispielsweise eine Suppression/Deletion oder Überexpression von HARP im Zusammenhang mit den RNA-Seq.-Analysen nicht möglich ist. Da jedoch viele der Zellen kultivierbar sind, könnte die Expression von HARP in Abhängigkeit unterschiedlicher Umweltfaktoren (z.B. unter verschiedenen Stressbedingungen) untersucht werden. Mit Hinblick auf die Evolution von RNase P ist das Bakterium *Thermodesulfatator indicus*, welches neben den Genen für die RNA-basierte RNase P ebenfalls ein HARP-Gen besitzt, von besonderem Interesse. Während das rekombinante HARP-Enzym eine hohe RNase P-Aktivität aufweist, konnte bei dem RNA-basierten Enzym nur eine niedrige Aktivität gemessen werden. Folglich könnte es sich hier um einen evolutiven Übergang vom RNA-basierten zum Protein-basierten Enzym handeln. Auch sollte die Struktur der HARP-Enzyme aufgeklärt werden, da viele Eigenschaften (z.B. Oligomerisierung, Substraterkennung/-bindung) dieser neuen Enzyme noch vollkommen unverstanden sind.

## 6 Literatur

1. Toh, Y., et al., *Transfer RNA synthesis and regulation*. eLS, 2009.
2. Altman, S. and J.D. Smith, *Tyrosine tRNA precursor molecule polynucleotide sequence*. Nat New Biol, 1971. **233**(36): p. 35-9.
3. Altman, S., *Biosynthesis of transfer RNA in Escherichia coli*. Cell, 1975. **4**(1): p. 21-9.
4. Kole, R. and S. Altman, *Reconstitution of RNase P activity from inactive RNA and protein*. Proc Natl Acad Sci U S A, 1979. **76**(8): p. 3795-9.
5. Sekiya, T., et al., *Total synthesis of a tyrosine suppressor transfer RNA gene. XVII. Transcription, in vitro, of the synthetic gene and processing of the primary transcript to transfer RNA*. J Biol Chem, 1979. **254**(13): p. 5802-16.
6. Mörl, M. and A. Marchfelder, *The final cut. The importance of tRNA 3'-processing*. EMBO Rep, 2001. **2**(1): p. 17-20.
7. Betat, H., et al., *From end to end: tRNA editing at 5'- and 3'-terminal positions*. Int J Mol Sci, 2014. **15**(12): p. 23975-98.
8. Ray, B.K. and D. Apirion, *RNAase P is dependent on RNAase E action in processing monomeric RNA precursors that accumulate in an RNAase E- mutant of Escherichia coli*. J Mol Biol, 1981. **149**(4): p. 599-617.
9. Ray, B.K. and D. Apirion, *Transfer RNA precursors are accumulated in Escherichia coli in the absence of RNase E*. Eur J Biochem, 1981. **114**(3): p. 517-24.
10. Apirion, D. and A. Miczak, *RNA processing in prokaryotic cells*. Bioessays, 1993. **15**(2): p. 113-20.
11. Li, Z. and M.P. Deutscher, *RNase E plays an essential role in the maturation of Escherichia coli tRNA precursors*. RNA, 2002. **8**(1): p. 97-109.
12. Lund, E. and J.E. Dahlberg, *Spacer transfer RNAs in ribosomal RNA transcripts of E. coli: processing of 30S ribosomal RNA in vitro*. Cell, 1977. **11**(2): p. 247-62.
13. Dunn, J.J., *RNase III cleavage of single-stranded RNA. Effect of ionic strength on the fidelity of cleavage*. J Biol Chem, 1976. **251**(12): p. 3807-14.
14. Ghosh, R.K. and M.P. Deutscher, *Identification of an Escherichia coli nuclease acting on structurally altered transfer RNA molecules*. J Biol Chem, 1978. **253**(4): p. 997-1000.
15. Cudny, H. and M.P. Deutscher, *Apparent involvement of ribonuclease D in the 3' processing of tRNA precursors*. Proc Natl Acad Sci U S A, 1980. **77**(2): p. 837-41.
16. Söll, D. and U.L. RajBhandary, *tRNA: Structure, Biosynthesis, and Function*. 1995.
17. El Yacoubi, B., M. Bailly, and V. de Crecy-Lagard, *Biosynthesis and function of posttranscriptional modifications of transfer RNAs*. Annu Rev Genet, 2012. **46**: p. 69-95.
18. Shepherd, J. and M. Ibba, *Bacterial transfer RNAs*. FEMS Microbiol Rev, 2015. **39**(3): p. 280-300.
19. Ericsson, U.B., P. Nordlund, and B.M. Hallberg, *X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold*. FEBS Lett, 2004. **565**(1-3): p. 59-64.

20. Hur, S. and R.M. Stroud, *How U38, 39, and 40 of many tRNAs become the targets for pseudouridylation by TruA*. Mol Cell, 2007. **26**(2): p. 189-203.
21. Ishida, K., et al., *Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium Thermus thermophilus*. Nucleic Acids Res, 2011. **39**(6): p. 2304-18.
22. Rintala-Dempsey, A.C. and U. Kothe, *Eukaryotic stand-alone pseudouridine synthases - RNA modifying enzymes and emerging regulators of gene expression?* RNA Biol, 2017. **14**(9): p. 1185-1196.
23. Winther, K.S. and K. Gerdes, *Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA*. Proc Natl Acad Sci U S A, 2011. **108**(18): p. 7403-7.
24. Robertson, H.D., S. Altman, and J.D. Smith, *Purification and properties of a specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor*. J Biol Chem, 1972. **247**(16): p. 5243-51.
25. Schedl, P. and P. Primakoff, *Mutants of Escherichia coli thermosensitive for the synthesis of transfer RNA*. Proc Natl Acad Sci U S A, 1973. **70**(7): p. 2091-5.
26. Stark, B.C., et al., *Ribonuclease P: an enzyme with an essential RNA component*. Proc Natl Acad Sci U S A, 1978. **75**(8): p. 3717-21.
27. Kole, R., et al., *E. coli RNAase P has a required RNA component*. Cell, 1980. **19**(4): p. 881-7.
28. Guerrier-Takada, C., et al., *The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme*. Cell, 1983. **35**(3 Pt 2): p. 849-57.
29. Guerrier-Takada, C. and S. Altman, *Structure in solution of M1 RNA, the catalytic subunit of ribonuclease P from Escherichia coli*. Biochemistry, 1984. **23**(26): p. 6327-34.
30. Reich, C., et al., *The RNA component of the Bacillus subtilis RNase P. Sequence, activity, and partial secondary structure*. J Biol Chem, 1986. **261**(17): p. 7888-93.
31. Burkard, U., I. Willis, and D. Soll, *Processing of histidine transfer RNA precursors. Abnormal cleavage site for RNase P*. J Biol Chem, 1988. **263**(5): p. 2447-51.
32. Reich, C., et al., *Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme*. Science, 1988. **239**(4836): p. 178-81.
33. Waugh, D.S. and N.R. Pace, *Complementation of an RNase P RNA (rnpB) gene deletion in Escherichia coli by homologous genes from distantly related eubacteria*. J Bacteriol, 1990. **172**(11): p. 6316-22.
34. Wegscheid, B., C. Condon, and R.K. Hartmann, *Type A and B RNase P RNAs are interchangeable in vivo despite substantial biophysical differences*. EMBO Rep, 2006. **7**(4): p. 411-7.
35. Gosringer, M. and R.K. Hartmann, *Function of heterologous and truncated RNase P proteins in Bacillus subtilis*. Mol Microbiol, 2007. **66**(3): p. 801-13.
36. Wegscheid, B. and R.K. Hartmann, *In vivo and in vitro investigation of bacterial type B RNase P interaction with tRNA 3'-CCA*. Nucleic Acids Res, 2007. **35**(6): p. 2060-73.

37. Pannucci, J.A., et al., *RNase P RNAs from some Archaea are catalytically active*. Proc Natl Acad Sci U S A, 1999. **96**(14): p. 7803-8.
38. Kikovska, E., S.G. Svard, and L.A. Kirsebom, *Eukaryotic RNase P RNA mediates cleavage in the absence of protein*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2062-7.
39. Doherty, E.A. and J.A. Doudna, *Ribozyme structures and mechanisms*. Annu Rev Biophys Biomol Struct, 2001. **30**: p. 457-75.
40. Blount, K.F. and O.C. Uhlenbeck, *The hammerhead ribozyme*. Biochem Soc Trans, 2002. **30**(Pt 6): p. 1119-22.
41. Rossmannith, W., et al., *Human mitochondrial tRNA processing*. J Biol Chem, 1995. **270**(21): p. 12885-91.
42. Rossmannith, W. and R.M. Karwan, *Characterization of human mitochondrial RNase P: novel aspects in tRNA processing*. Biochem Biophys Res Commun, 1998. **247**(2): p. 234-41.
43. Thomas, B.C., et al., *Spinach chloroplast RNase P: a putative protein enzyme*. Nucleic Acids Symp Ser, 1995(33): p. 95-8.
44. Wang, M.J., N.W. Davis, and P. Gegenheimer, *Novel mechanisms for maturation of chloroplast transfer RNA precursors*. EMBO J, 1988. **7**(6): p. 1567-74.
45. Holzmann, J., et al., *RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme*. Cell, 2008. **135**(3): p. 462-74.
46. Gobert, A., et al., *A single Arabidopsis organellar protein has RNase P activity*. Nat Struct Mol Biol, 2010. **17**(6): p. 740-4.
47. Gutmann, B., A. Gobert, and P. Giege, *PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis*. Genes Dev, 2012. **26**(10): p. 1022-7.
48. Taschner, A., et al., *Nuclear RNase P of Trypanosoma brucei: a single protein in place of the multicomponent RNA-protein complex*. Cell Rep, 2012. **2**(1): p. 19-25.
49. Nickel, A.I., et al., *Minimal and RNA-free RNase P in Aquifex aeolicus*. Proc Natl Acad Sci U S A, 2017. **114**(42): p. 11121-11126.
50. Karasik, A., et al., *Nuclear Protein-Only Ribonuclease P2 Structure and Biochemical Characterization Provide Insight into the Conserved Properties of tRNA 5' End Processing Enzymes*. J Mol Biol, 2016. **428**(1): p. 26-40.
51. Howard, M.J., et al., *Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing*. Proc Natl Acad Sci U S A, 2012. **109**(40): p. 16149-54.
52. Reinhard, L., S. Sridhara, and B.M. Hallberg, *Structure of the nuclease subunit of human mitochondrial RNase P*. Nucleic Acids Res, 2015. **43**(11): p. 5664-72.
53. Ozeki, H., et al., *Temperature-sensitive mutants of Escherichia coli defective in tRNA biosynthesis*. Brookhaven Symp Biol, 1975(26): p. 89-105.
54. Haas, E.S., et al., *Further perspective on the catalytic core and secondary structure of ribonuclease P RNA*. Proc Natl Acad Sci U S A, 1994. **91**(7): p. 2527-31.
55. Harris, J.K., et al., *New insight into RNase P RNA structure from comparative analysis of the archaeal RNA*. RNA, 2001. **7**(2): p. 220-32.

56. Brown, J.W., et al., *Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements*. Proc Natl Acad Sci U S A, 1996. **93**(7): p. 3001-6.
57. Haas, E.S., et al., *Structure and evolution of ribonuclease P RNA in Gram-positive bacteria*. Nucleic Acids Res, 1996. **24**(23): p. 4775-82.
58. Brown, J.W. and N.R. Pace, *Ribonuclease P RNA and protein subunits from bacteria*. Nucleic Acids Res, 1992. **20**(7): p. 1451-6.
59. Kazantsev, A.V. and N.R. Pace, *Bacterial RNase P: a new view of an ancient enzyme*. Nat Rev Microbiol, 2006. **4**(10): p. 729-40.
60. Hartmann, R.K., et al., *The making of tRNAs and more - RNase P and tRNase Z*. Prog Mol Biol Transl Sci, 2009. **85**: p. 319-68.
61. Evans, D., S.M. Marquez, and N.R. Pace, *RNase P: interface of the RNA and protein worlds*. Trends Biochem Sci, 2006. **31**(6): p. 333-41.
62. Westhof, E., D. Wesolowski, and S. Altman, *Mapping in three dimensions of regions in a catalytic RNA protected from attack by an Fe(II)-EDTA reagent*. J Mol Biol, 1996. **258**(4): p. 600-13.
63. Kim, J.J., et al., *The protein cofactor allows the sequence of an RNase P ribozyme to diversify by maintaining the catalytically active structure of the enzyme*. RNA, 1997. **3**(6): p. 613-23.
64. Buck, A.H., et al., *Protein activation of a ribozyme: the role of bacterial RNase P protein*. EMBO J, 2005. **24**(19): p. 3360-8.
65. Buck, A.H., et al., *Structural perspective on the activation of RNase P RNA by protein*. Nat Struct Mol Biol, 2005. **12**(11): p. 958-64.
66. Crary, S.M., S. Niranjana, and C.A. Fierke, *The protein component of Bacillus subtilis ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNA<sup>Asp</sup>*. Biochemistry, 1998. **37**(26): p. 9409-16.
67. Kurz, J.C., S. Niranjana, and C.A. Fierke, *Protein component of Bacillus subtilis RNase P specifically enhances the affinity for precursor-tRNA<sup>Asp</sup>*. Biochemistry, 1998. **37**(8): p. 2393-400.
68. Niranjana, S., J.C. Kurz, and C.A. Fierke, *Expression, purification and characterization of the recombinant ribonuclease P protein component from Bacillus subtilis*. Nucleic Acids Res, 1998. **26**(13): p. 3090-6.
69. Guo, X., et al., *RNA-dependent folding and stabilization of C5 protein during assembly of the E. coli RNase P holoenzyme*. J Mol Biol, 2006. **360**(1): p. 190-203.
70. Spitzfaden, C., et al., *The structure of ribonuclease P protein from Staphylococcus aureus reveals a unique binding site for single-stranded RNA*. J Mol Biol, 2000. **295**(1): p. 105-15.
71. Kazantsev, A.V., et al., *High-resolution structure of RNase P protein from Thermotoga maritima*. Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7497-502.
72. Stams, T., et al., *Ribonuclease P protein structure: evolutionary origins in the translational apparatus*. Science, 1998. **280**(5364): p. 752-5.
73. Gößringer, M. and R.K. Hartmann, *Function of heterologous and truncated RNase P proteins in Bacillus subtilis*. Mol Microbiol, 2007. **66**(3): p. 801-13.

74. Jarrous, N. and V. Gopalan, *Archaeal/eukaryal RNase P: subunits, functions and RNA diversification*. Nucleic Acids Res, 2010. **38**(22): p. 7885-94.
75. Marvin, M.C. and D.R. Engelke, *RNase P: increased versatility through protein complexity?* RNA Biol, 2009. **6**(1): p. 40-2.
76. Wan, F., et al., *Cryo-electron microscopy structure of an archaeal ribonuclease P holoenzyme*. Nat Commun, 2019. **10**(1): p. 2617.
77. Lan, P., et al., *Structural insight into precursor tRNA processing by yeast ribonuclease P*. Science, 2018. **362**(6415).
78. Jackman, J.E., et al., *Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9*. RNA, 2003. **9**(5): p. 574-85.
79. Jörnvall, H., et al., *Short-chain dehydrogenases/reductases (SDR)*. Biochemistry, 1995. **34**(18): p. 6003-13.
80. Lechner, M., et al., *Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya*. Mol Biol Evol, 2015. **32**(12): p. 3186-93.
81. Sugita, C., et al., *Molecular characterization of three PRORP proteins in the moss Physcomitrella patens: nuclear PRORP protein is not essential for moss viability*. PLoS One, 2014. **9**(10): p. e108962.
82. Anantharaman, V. and L. Aravind, *The NYN domains: novel predicted RNases with a PIN domain-like fold*. RNA Biol, 2006. **3**(1): p. 18-27.
83. Schmitz-Linneweber, C. and I. Small, *Pentatricopeptide repeat proteins: a socket set for organelle gene expression*. Trends Plant Sci, 2008. **13**(12): p. 663-70.
84. Lurin, C., et al., *Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis*. Plant Cell, 2004. **16**(8): p. 2089-103.
85. Oerum, S., et al., *Crystal structure of the methyltransferase subunit of human mitochondrial Ribonuclease P (MRPP1) bound to S-adenosyl-methionine (SAM)*. to be published.
86. Kavanagh, K.L., et al., *The structure of wild-type human HADH2 (17beta-hydroxysteroid dehydrogenase type 10) bound to NAD+ at 1.2 Å*. to be published.
87. Walczyk, D., et al., *Analysis of the Cleavage Mechanism by Protein-Only RNase P Using Precursor tRNA Substrates with Modifications at the Cleavage Site*. J Mol Biol, 2016. **428**(24 Pt B): p. 4917-4928.
88. Smith, D., et al., *Influence of metal ions on the ribonuclease P reaction. Distinguishing substrate binding from catalysis*. J Biol Chem, 1992. **267**(4): p. 2429-36.
89. Sprinzl, M., et al., *Compilation of tRNA sequences and sequences of tRNA genes*. Nucleic Acids Res, 1996. **24**(1): p. 68-72.
90. Christian, E.L., et al., *The P4 metal binding site in RNase P RNA affects active site metal affinity through substrate positioning*. RNA, 2006. **12**(8): p. 1463-7.
91. Holm, P.S. and G. Krupp, *The acceptor stem in pre-tRNAs determines the cleavage specificity of RNase P*. Nucleic Acids Res, 1992. **20**(3): p. 421-3.
92. Kirsebom, L.A. and S.G. Svard, *The kinetics and specificity of cleavage by RNase P is mainly dependent on the structure of the amino acid acceptor stem*. Nucleic Acids Res, 1992. **20**(3): p. 425-32.



93. Sun, L., et al., *Evidence that substrate-specific effects of C5 protein lead to uniformity in binding and catalysis by RNase P*. EMBO J, 2006. **25**(17): p. 3998-4007.
94. Koutmou, K.S., et al., *Protein-precursor tRNA contact leads to sequence-specific recognition of 5' leaders by bacterial ribonuclease P*. J Mol Biol, 2010. **396**(1): p. 195-208.
95. Niranjanakumari, S., et al., *Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA*. Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15212-7.
96. Rueda, D., et al., *The 5' leader of precursor tRNA<sup>Asp</sup> bound to the Bacillus subtilis RNase P holoenzyme has an extended conformation*. Biochemistry, 2005. **44**(49): p. 16130-9.
97. Brännvall, M., et al., *Evidence for induced fit in bacterial RNase P RNA-mediated cleavage*. J Mol Biol, 2007. **372**(5): p. 1149-64.
98. Pan, T., A. Loria, and K. Zhong, *Probing of tertiary interactions in RNA: 2'-hydroxyl-base contacts between the RNase P RNA and pre-tRNA*. Proc Natl Acad Sci U S A, 1995. **92**(26): p. 12510-4.
99. Loria, A. and T. Pan, *Recognition of the T stem-loop of a pre-tRNA substrate by the ribozyme from Bacillus subtilis ribonuclease P*. Biochemistry, 1997. **36**(21): p. 6317-25.
100. McClain, W.H., C. Guerrier-Takada, and S. Altman, *Model substrates for an RNA enzyme*. Science, 1987. **238**(4826): p. 527-30.
101. Kirsebom, L.A. and S.G. Svard, *Base pairing between Escherichia coli RNase P RNA and its substrate*. EMBO J, 1994. **13**(20): p. 4870-6.
102. Kahle, D., et al., *The methylation of one specific guanosine in a pre-tRNA prevents cleavage by RNase P and by the catalytic M1 RNA*. Nucleic Acids Res, 1990. **18**(4): p. 837-44.
103. Kahle, D., U. Wehmeyer, and G. Krupp, *Substrate recognition by RNase P and by the catalytic M1 RNA: identification of possible contact points in pre-tRNAs*. EMBO J, 1990. **9**(6): p. 1929-37.
104. Sinapah, S., et al., *Cleavage of model substrates by archaeal RNase P: role of protein cofactors in cleavage-site selection*. Nucleic Acids Res, 2011. **39**(3): p. 1105-16.
105. McClain, W.H. and H.B. Nicholas, Jr., *Differences between transfer RNA molecules*. J Mol Biol, 1987. **194**(4): p. 635-42.
106. Harris, M.E. and E.L. Christian, *Recent insights into the structure and function of the ribonucleoprotein enzyme ribonuclease P*. Curr Opin Struct Biol, 2003. **13**(3): p. 325-33.
107. Howard, M.J., et al., *Differential substrate recognition by isozymes of plant protein-only Ribonuclease P*. RNA, 2016. **22**(5): p. 782-92.
108. Brillante, N., et al., *Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P*. Nucleic Acids Res, 2016. **44**(5): p. 2323-36.
109. Klemm, B.P., et al., *Molecular recognition of pre-tRNA by Arabidopsis protein-only Ribonuclease P*. RNA, 2017. **23**(12): p. 1860-1873.
110. Gobert, A., et al., *Structural insights into protein-only RNase P complexed with tRNA*. Nat Commun, 2013. **4**: p. 1353.
111. Imai, T., et al., *Pentatricopeptide repeat motifs in the processing enzyme PRORP1 in Arabidopsis thaliana play a crucial role in recognition of*

- nucleotide bases at *TpsiC* loop in precursor tRNAs. *Biochem Biophys Res Commun*, 2014. **450**(4): p. 1541-6.
112. Sambrook, J., E.F. Fritsch, and T.E. Maniatis, *Molecular cloning. A laboratory manual*. Cold spring Harbor Laboratory, Cold Spring Harbor, N.Y, 1989.
113. Wegscheid, B. and R.K. Hartmann, *The precursor tRNA 3'-CCA interaction with Escherichia coli RNase P RNA is essential for catalysis by RNase P in vivo*. *RNA*, 2006. **12**(12): p. 2135-48.
114. Gossringer, M., R. Kretschmer-Kazemi Far, and R.K. Hartmann, *Analysis of RNase P protein (rnpA) expression in Bacillus subtilis utilizing strains with suppressible rnpA expression*. *J Bacteriol*, 2006. **188**(19): p. 6816-23.
115. Vehmaanpera, J., *Transformation of Bacillus amyloliquefaciens by electroporation*. *FEMS Microbiol Lett*, 1989. **52**(1-2): p. 165-9.
116. Hanahan, D., *Studies on transformation of Escherichia coli with plasmids*. *J Mol Biol*, 1983. **166**(4): p. 557-80.
117. Kirby, K.S., *A new method for the isolation of ribonucleic acids from mammalian tissues*. *Biochem J*, 1956. **64**(3): p. 405-8.
118. Shapiro, D.J., *Quantitative ethanol precipitation of nanogram quantities of DNA and RNA*. *Anal Biochem*, 1981. **110**(1): p. 229-31.
119. Studier, F.W., *Protein production by auto-induction in high density shaking cultures*. *Protein Expr Purif*, 2005. **41**(1): p. 207-34.
120. Grodberg, J. and J.J. Dunn, *ompT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification*. *J Bacteriol*, 1988. **170**(3): p. 1245-53.
121. Zawadzki, V. and H.J. Gross, *Rapid and simple purification of T7 RNA polymerase*. *Nucleic Acids Res*, 1991. **19**(8): p. 1948.
122. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. *Anal Biochem*, 1976. **72**: p. 248-54.
123. Donaldson, G.P., et al., *A rapid assay for affinity and kinetics of molecular interactions with nucleic acids*. *Nucleic Acids Res*, 2012. **40**(7): p. e48.
124. Fadouloglou, V.E., M. Kokkinidis, and N.M. Glykos, *Determination of protein oligomerization state: two approaches based on glutaraldehyde crosslinking*. *Anal Biochem*, 2008. **373**(2): p. 404-6.
125. Niland, C.N., et al., *Determination of the Specificity Landscape for Ribonuclease P Processing of Precursor tRNA 5' Leader Sequences*. *ACS Chem Biol*, 2016. **11**(8): p. 2285-92.
126. Niland, C.N., et al., *The contribution of the C5 protein subunit of Escherichia coli ribonuclease P to specificity for precursor tRNA is modulated by proximal 5' leader sequences*. *RNA*, 2017. **23**(10): p. 1502-1511.
127. Svärd, S.G., U. Kagardt, and L.A. Kirsebom, *Phylogenetic comparative mutational analysis of the base-pairing between RNase P RNA and its substrate*. *RNA*, 1996. **2**(5): p. 463-72.
128. Hardt, W.D., et al., *Role of the D arm and the anticodon arm in tRNA recognition by eubacterial and eukaryotic RNase P enzymes*. *Biochemistry*, 1993. **32**(48): p. 13046-53.

129. Wu, S., et al., *Cleavage mediated by the catalytic domain of bacterial RNase P RNA*. J Mol Biol, 2012. **422**(2): p. 204-14.
130. Chan, C.W., B. Chetnani, and A. Mondragon, *Structure and function of the T-loop structural motif in noncoding RNAs*. Wiley Interdiscip Rev RNA, 2013. **4**(5): p. 507-22.
131. Giege, R., et al., *Structure of transfer RNAs: similarity and variability*. Wiley Interdiscip Rev RNA, 2012. **3**(1): p. 37-61.
132. Hardt, W.D., et al., *Kinetics and thermodynamics of the RNase P RNA cleavage reaction: analysis of tRNA 3'-end variants*. J Mol Biol, 1995. **247**(2): p. 161-72.
133. Challacombe, J.F., et al., *Complete genome sequence of Halorhodospira halophila SL1*. Stand Genomic Sci, 2013. **8**(2): p. 206-14.
134. Hoeft, S.E., et al., *Alkalilimnicola ehrlichii sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor*. Int J Syst Evol Microbiol, 2007. **57**(Pt 3): p. 504-12.
135. Tikhonova, T.V., et al., *Molecular and catalytic properties of a novel cytochrome c nitrite reductase from nitrate-reducing haloalkaliphilic sulfur-oxidizing bacterium Thioalkalivibrio nitratireducens*. Biochim Biophys Acta, 2006. **1764**(4): p. 715-23.
136. Hou, S., et al., *Complete genome sequence of the extremely acidophilic methanotroph isolate V4, Methylacidiphilum infernorum, a representative of the bacterial phylum Verrucomicrobia*. Biol Direct, 2008. **3**: p. 26.
137. Moussard, H., et al., *Thermodesulfatator indicus gen. nov., sp. nov., a novel thermophilic chemolithoautotrophic sulfate-reducing bacterium isolated from the Central Indian Ridge*. Int J Syst Evol Microbiol, 2004. **54**(Pt 1): p. 227-33.
138. Saxowsky, T.T., et al., *Trypanosoma brucei has two distinct mitochondrial DNA polymerase beta enzymes*. J Biol Chem, 2003. **278**(49): p. 49095-101.
139. Raes, J. and Y. Van de Peer, *Gene duplication, the evolution of novel gene functions, and detecting functional divergence of duplicates in silico*. Appl Bioinformatics, 2003. **2**(2): p. 91-101.
140. Bielawski, J.P. and Z. Yang, *Maximum likelihood methods for detecting adaptive evolution after gene duplication*. J Struct Funct Genomics, 2003. **3**(1-4): p. 201-12.
141. Qian, W., et al., *Maintenance of duplicate genes and their functional redundancy by reduced expression*. Trends Genet, 2010. **26**(10): p. 425-30.
142. Koonin, E.V., K.S. Makarova, and L. Aravind, *Horizontal gene transfer in prokaryotes: quantification and classification*. Annu Rev Microbiol, 2001. **55**: p. 709-42.
143. Koonin, E.V., *Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions*. F1000Res, 2016. **5**.
144. Aravind, L., et al., *Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles*. Trends Genet, 1998. **14**(11): p. 442-4.
145. Ohno, S., *Evolution by Gene Duplication*. 1970: Springer Verlag.

146. Schwarz, T.S., et al., *Homologs of aquifex aeolicus protein-only RNase P are not the major RNase P activities in the archaea haloferax volcanii and methanosarcina mazei*. IUBMB Life, 2019. **71**(8): p. 1109-1116.
147. Gobert, A., M. Bruggeman, and P. Giege, *Involvement of PIN-like domain nucleases in tRNA processing and translation regulation*. IUBMB Life, 2019.
148. Krieger, E., S.B. Nabuurs, and G. Vriend, *Homology modeling*. Methods Biochem Anal, 2003. **44**: p. 509-23.
149. Xiang, Z., *Advances in homology protein structure modeling*. Curr Protein Pept Sci, 2006. **7**(3): p. 217-27.
150. Andrews, P., *Estimation of the molecular weights of proteins by Sephadex gel-filtration*. Biochem J, 1964. **91**(2): p. 222-33.
151. Prunetti, L., et al., *New functional sulfide oxidase-oxygen reductase supercomplex in the membrane of the hyperthermophilic bacterium Aquifex aeolicus*. J Biol Chem, 2010. **285**(53): p. 41815-26.
152. Guiral, M., et al., *The hyperthermophilic bacterium Aquifex aeolicus: from respiratory pathways to extremely resistant enzymes and biotechnological applications*. Adv Microb Physiol, 2012. **61**: p. 125-94.
153. Urbanke, C., G. Witte, and U. Curth, *Sedimentation velocity method in the analytical ultracentrifuge for the study of protein-protein interactions*. Methods Mol Biol, 2005. **305**: p. 101-14.
154. Palleros, D.R., et al., *hsp70-protein complexes. Complex stability and conformation of bound substrate protein*. J Biol Chem, 1994. **269**(18): p. 13107-14.
155. de Lichtenberg, U., et al., *Dynamic complex formation during the yeast cell cycle*. Science, 2005. **307**(5710): p. 724-7.
156. Wahle, E., R.S. Lasken, and A. Kornberg, *The dnaB-dnaC replication protein complex of Escherichia coli. I. Formation and properties*. J Biol Chem, 1989. **264**(5): p. 2463-8.
157. Ferre-D'Amare, A.R. and S.K. Burley, *Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies*. Structure, 1994. **2**(5): p. 357-9.
158. Ranjbar, B. and P. Gill, *Circular dichroism techniques: biomolecular and nanostructural analyses- a review*. Chem Biol Drug Des, 2009. **74**(2): p. 101-20.
159. Kwan, A.H., et al., *Macromolecular NMR spectroscopy for the non-spectroscopist*. FEBS J, 2011. **278**(5): p. 687-703.
160. Chernov, A.A., *Protein crystals and their growth*. J Struct Biol, 2003. **142**(1): p. 3-21.
161. Carter, C.W., Jr., *[5] Response surface methods for optimizing and improving reproducibility of crystal growth*. Methods Enzymol, 1997. **276**: p. 74-99.
162. Schmit, J.D. and K.A. Dill, *The stabilities of protein crystals*. J Phys Chem B, 2010. **114**(11): p. 4020-7.
163. Schmit, J.D. and K. Dill, *Growth rates of protein crystals*. J Am Chem Soc, 2012. **134**(9): p. 3934-7.
164. Bergfors, T., *Seeds to crystals*. J Struct Biol, 2003. **142**(1): p. 66-76.
165. Hallberg, B.M., et al., *A new scaffold for binding haem in the cytochrome domain of the extracellular flavocytochrome cellobiose dehydrogenase*. Structure, 2000. **8**(1): p. 79-88.

166. Sousa, R., et al., *Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution*. Nature, 1993. **364**(6438): p. 593-9.
167. Sousa, R., *Structural and mechanistic relationships between nucleic acid polymerases*. Trends Biochem Sci, 1996. **21**(5): p. 186-90.
168. Sousa, R., *Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization*. Acta Crystallogr D Biol Crystallogr, 1995. **51**(Pt 3): p. 271-7.

## Anhang

## BioSpektrum – Artikel 2018

582 WISSENSCHAFT

## Evolution der tRNA-Prozessierung

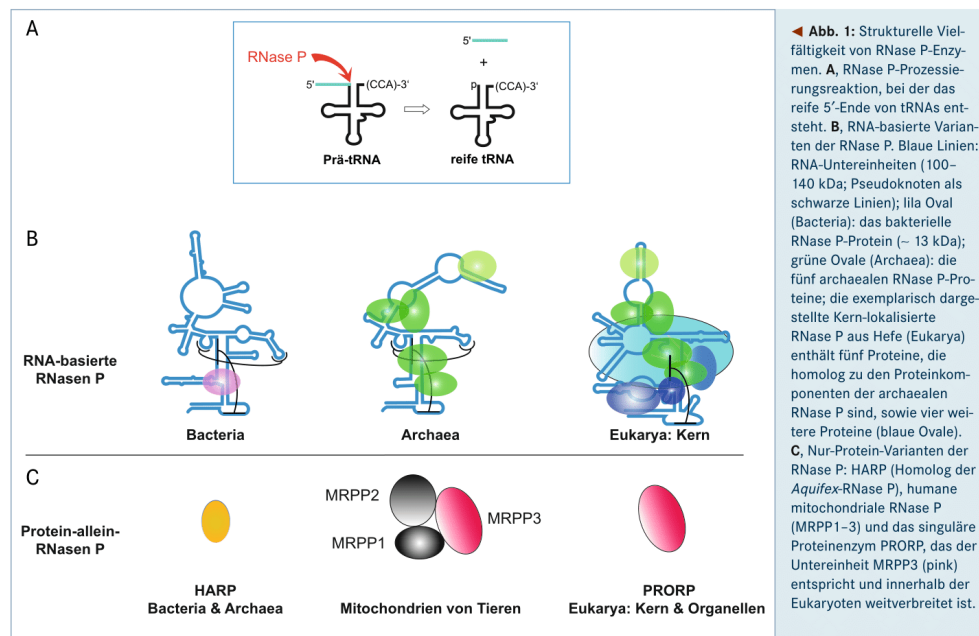
## Minimale RNA-freie RNase P in einem hyperthermophilen Bakterium

ASTRID I. NICKEL, NADINE B. WÄBER, MARKUS GÖBRINGER,  
ISABELL SCHENCKING, ROLAND K. HARTMANN  
INSTITUT FÜR PHARMAZEUTISCHE CHEMIE, UNIVERSITÄT MARBURG

The essential tRNA processing enzyme RNase P consists of a catalytic RNA and a small protein in most Bacteria. An RNA was also identified in archaeal and many eukaryal nuclear RNases P, with five protein cofactors in Archaea and nine or ten in Eukarya. Yet, more recently, the second RNase P in human cells, acting in mitochondria, was found to lack an RNA; this type of protein-only RNase P is also wide-spread among eukaryotes. Moreover, we recently discovered a novel RNA-free RNase P in the hyperthermophilic bacterium *Aquifex aeolicus*, consisting of a single polypeptide of 23 kilodaltons.

DOI: 10.1007/s12268-018-0960-3  
© Springer-Verlag 2018

■ Bisher ging man davon aus, dass das essenzielle tRNA-Prozessierungsenzym RNase P (**Abb. 1**) in Bakterien generell aus einer katalytischen RNA und einem kleinen Protein besteht. Eine RNA-Komponente ist auch in RNase P-Enzymen aus Archaea sowie vielen eukaryotischen Kernen enthalten, die jedoch mit mehr Proteinuntereinheiten (fünf in Archaeen, neun bis zehn in Eukaryoten) assoziiert ist (**Abb. 1B**). 2008 fand die Wiener Arbeitsgruppe um Walter Rossmanith heraus, dass die humane mitochondriale RNase P – die zweite RNase P neben dem im Kern lokalisierten, RNA-basierten Enzym – keine RNA-Komponente besitzt [1]. Diese Art der rein proteinbasierten RNase P (MRPP3, PRORP; **Abb. 1C**), die man nicht in Bakterien und Archaeen



findet, ist innerhalb der Eukaryoten weit verbreitet. Doch auch in Bakterien gibt es Ausnahmen von der Regel: So fanden wir im hyperthermophilen Bakterium *Aquifex aeolicus* eine RNA-freie RNase P (HARP; **Abb. 1C**), die nur aus einem 23 Kilodalton großen Polypeptid besteht [2].

#### Der Fall *Aquifex aeolicus*

Auf dem Gebiet der bakteriellen tRNA-Prozessierung gab es bis 2017 einen letzten „ungelösten Fall“, den 1998 die Veröffentlichung der Genomsequenz des hyperthermophilen Bakteriums *Aquifex aeolicus* (Familie der Aquificaceae) ausgelöst hatte [3]. So scheiterten sämtliche bioinformatischen Anstrengungen verschiedener Arbeitsgruppen, die Gene für die RNA- (*rnpB*) und Proteinuntereinheit (*rnpA*) der RNase P in dem stark kondensierten Genom (1,55 Megabasenpaare) zu identifizieren. Auch gelang es uns zunächst nicht, RNase-P-Aktivität in partiell gereinigten Zellextrakten von *A. aeolicus* nachzuweisen [4]. Wir fanden schließlich heraus, dass

eine robuste RNase-P-Aktivität in Zellextrakten von *A. aeolicus* nachweisbar wurde, wenn die dabei verwendeten Puffer die nicht-flüchtige Thiol-reduzierende Verbindung Dithiothreitol (DTT) enthielten [5]. Der Nachweis der Enzymaktivität ermöglichte in der Folge gezielte biochemische Experimente zum Aufbau des Enzyms in *Aquifex*. Dabei sahen wir [5], im Gegensatz zu einer anderen Arbeitsgruppe [6], keine Evidenz für das Vorhandensein einer RNA-Untereinheit. Unsere Schlussfolgerung basierte auf der Beobachtung, dass das Enzym durch Behandlung mit Mikrokokken-Nuklease, die eine mögliche RNA-Komponente abbauen sollte, nicht inaktiviert wurde.

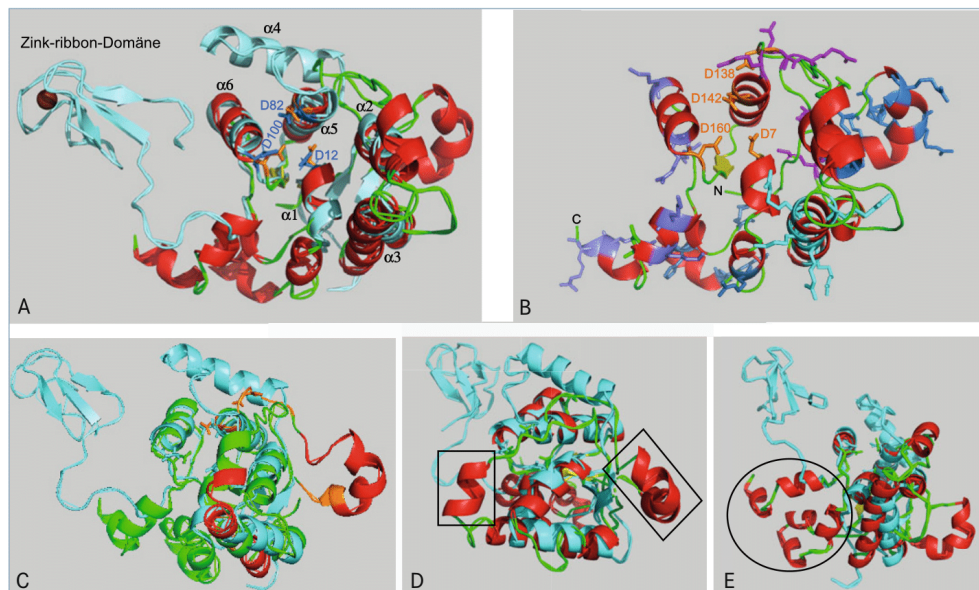
#### Biochemische Reinigung und Identifizierung der RNase P

Da die Zellanzucht des hyperthermophilen Bakteriums *A. aeolicus* (Wachstum bei 67–95 °C) unter mikroaeroben Bedingungen eine besondere Laborausstattung erfordert, nahmen wir den Fermentationsservice des

Archaeenzentrums der Universität Regensburg in Anspruch. Astrid I. Nickel erarbeitete im Rahmen ihrer Doktorarbeit ein Protokoll zur Anreicherung der *A. aeolicus*-RNase P in drei chromatographischen Schritten (Anionenaustausch-, hydrophobe Interaktions- und Größenausschlusschromatographie). Säulenfraktionen der letzten beiden Chromatographiestritte analysierten wir mittels SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE), wobei hier die Verwendung von Stufengradientgelen wichtig war, um gleichzeitig große und kleine Proteine zu erfassen.

Proteinbanden, die in Säulenfraktionen mit angereicherter RNase P-Aktivität ebenfalls angereichert waren, identifizierten wir massenspektroskopisch. Aus einem halben Dutzend engerer Kandidatenproteine stach das 23 Kilodalton große Aq\_880 hervor, da es als einziges Protein sowohl in Fraktionen mit maximaler RNase P-Aktivität aus der Größenausschlusschromatographie als auch in der kleinsten von drei Aktivitätspeaks aus





▲ **Abb. 2:** Strukturmodell von Aq\_880. **A**, 3D-Homologiemodell von Aq\_880 auf Basis der NMR-Struktur von PhNob1 (mit Phyre<sup>2</sup> berechnet, mit STAMP strukturell aligniert). PhNob1 in Cyan, Aq\_880 in Rot ( $\alpha$ -Helices), Grün (flexible Bereiche) und Gelb ( $\beta$ -Stränge); die Zink-Ribbon-Domäne (Zn<sup>2+</sup>-Ion, braune Kugel) von PhNob1 fehlt in Aq\_880. Das Modell überlagert vier  $\alpha$ -Helices ( $\alpha$ 1,  $\alpha$ 2,  $\alpha$ 5 und  $\alpha$ 6) sowie eine fünfte ( $\alpha$ 3) teilweise. Die Asparate D12, D82 und D100, die ein einzelnes katalytisches Mn<sup>2+</sup>-Ion in PhNob1 koordinieren, sind blau markiert; diesen entsprechen D7, D142 und D160 in Aq\_880 (orange markiert in B). **B**, Häufungsregionen positiv geladener Seitenketten (Arg-, Lys- sowie zwei His-Reste) in Aq\_880, anhand verschiedener Blauschattierungen illustriert. N, C: N- und C-Terminus von Aq\_880. Die experimentell verifizierten, katalytisch relevanten Asparate D7, D138, D142, D160 (orange) koordinieren sehr wahrscheinlich ein bis zwei Mg<sup>2+</sup>- oder Mn<sup>2+</sup>-Ionen. **C**, Vorhersage von *coiled-coil*-Domänen-Interaktionen in Aq\_880 (rot und orange: hohe und mittlere Wahrscheinlichkeit); alle anderen Aminosäuren von Aq\_880 in Grün, von PhNob1 in Cyan. **D, E**, Zwei unterschiedliche Orientierungen, um die Aq\_880-spezifischen  $\alpha$ -Helices hervorzuheben (eingerahmt), die in PhNob1 nicht vorhanden sind.

der hydrophoben Interaktionschromatographie nachweisbar war. Rekombinant in *Escherichia coli* hergestelltes Aq\_880, für das eine Kontamination mit RNA-basierter *E. coli*-RNase P mittels RT-PCR ausgeschlossen werden kann, zeigt eine deutliche RNase P-Aktivität in Gegenwart von Mg<sup>2+</sup>- oder Mn<sup>2+</sup>-Ionen als Kofaktoren [2]. Aq\_880 ist wie erwartet hitzeresistent, und die katalytische Effizienz liegt im Bereich derjenigen anderer RNase P-Enzyme. Genetische Komplementationsanalysen bestätigen zudem die RNase P-Funktion. So wachsen sowohl *E. coli*- als auch Hefe-Zellen mit inaktivierter endogener RNase P wieder, wenn sie das Aq\_880-Protein exprimieren [2].

#### Aq\_880-Homologe in anderen Organismen

Neben homologen Proteinen in den bekannten Genomen anderer Bakterien der Familie der Aquificaceae identifizierten wir etwa 75 Homologe der *Aquifex*-RNase P (HARPs) in Archaeen sowie etwa zehn Homologe in anderen Bakterien, jedoch keine in Eukaryoten.

Das breite Vorkommen von HARPs bei Archaeen sowie die Erkenntnis aus bioinformatischen Genomvergleichen, dass *A. aeolicus* mindestens zehn Prozent seiner Protein-codierenden Gene durch horizontalen Gentransfer aus Archaeen akquiriert hat, machen es sehr wahrscheinlich, dass auch Aq\_880 von einem Archaeon in das Genom des Vorläuferbakteriums der Aquificaceae übertragen wurde.

Überraschend war weiterhin, dass alle Archaeen sowie die meisten anderen Bakterien außerhalb der Aquificaceae, die ein HARP codieren, nach wie vor die Gene für die klassische RNA-basierte RNase P besitzen. Wir konnten bisher für ein Archaeon (*Methanothermobacter thermautotrophicus*) und ein Bakterium (*Thermodesulfatator indicus*) nachweisen, dass neben der RNA-basierten RNase P auch das entsprechende HARP eine RNase-P-Aktivität besitzt.

Das Vorkommen zweier strukturell verschiedener Enzyme mit RNase-P-Aktivität im gleichen Organismus wirft verschiedene Fragen auf: Üben die beiden Enzyme die gleiche zelluläre Funktion aus, prozessieren sie also

beide sämtliche tRNA-Vorläufertranskripte? Oder arbeiten die beiden Enzyme arbeitsteilig, das heißt, prozessieren sie (schwerpunktmäßig) unterschiedliche Subgruppen von tRNA-Transkripten? Oder wird das HARP-Enzym unter bestimmten (Extrem-)Bedingungen verstärkt exprimiert, unter denen das RNA-basierte Enzym nicht mehr hergestellt wird oder inaktiv ist? Gibt es vielleicht eine Funktion von HARP, die über die klassische tRNA-Prozessierung hinausgeht?

#### HARPs: Oligomerisierung und RNA-Bindung

In der Größenausschlusschromatographie eluiert das rekombinante Aq\_880 schwermäßig bei 420 Kilodalton [2]. Diese Größe lässt sich durch die Bildung eines Trimers von Hexameren oder eines Hexamers aus Trimeren erklären. Welche Strukturelemente des Proteins die Oligomerisierung vermitteln sowie die Art der Wechselwirkung (z. B. *coiled-coil*-Domänen-Interaktion), untersuchen wir derzeit. Unklar ist auch, wie HARPs überhaupt RNA binden können bzw. spezifisch

tRNA-Transkripte erkennen und prozessieren. Eine RNA-Bindungsdomäne, wie man sie z. B. in den wesentlich größeren eukaryotischen *protein-only RNases P* (PRORPs, ~ 65 Kilodalton) findet, ist in den HARPs nicht erkennbar.

Eine andere Metallonuklease, Nob1, wird wie HARP der Pin\_5-Gruppe der Superfamilie der *Pin-like domain*-Metallonukleasen zugeordnet [7]. Nob1 ist in Archaeen und Eukaryoten an der Reifung der ribosomalen 16/18S-rRNA beteiligt [8]. Eine Homologie-modellierung von Aq\_880 auf Basis einer NMR-Struktur von Nob1 aus *Pyrococcus horikoshii* (PhNob1) [8] zeigt die überlappende Positionierung von drei katalytischen Asparataten sowie von fünf  $\alpha$ -Helices (Abb. 2A). Den vorhergesagten, weitestgehend  $\alpha$ -helikalen Charakter von Aq\_880 bestätigen CD-Spektren.

Nob1 besitzt eine zusätzliche Zink-Ribbon-Domäne, die die Spezifität der Bindung an eine Helix der 16/18S-rRNA vermittelt, während die isolierte Pin\_5-Domäne nur unspezifisch an einzelsträngige RNA bindet [8]. Diese Zink-Ribbon-Domäne fehlt in den HARPs, was ihre molekulare Basis der RNA-Erkennung weiterhin im Dunkeln lässt. Auf der Oberfläche des modellierten Aq\_880 gibt es einige Bereiche, in denen Aminosäuren mit positiv geladenen Seitenketten gehäuft vorkommen (Abb. 2B), die also RNA-Interaktionen vermitteln könnten. Zur Oligomerisierung von Aq\_880 sind bioinformatische Vorhersagen für mögliche *coiled-coil*-Domänen-Interaktionen, wie man sie z. B. von Transkriptionsfaktoren kennt, in Abbildung 2C dargestellt. Einige  $\alpha$ -Helices, die Aq\_880 und andere HARPs besitzen, die aber in PhNob1 nicht vorhanden sind (Abb. 2D, E), könnten Kandidaten sowohl für Oligomerisierungs- als auch Protein-RNA-Interaktionen sein.

#### Warum haben die Aquificaceae ihre RNA-basierte RNase P aufgegeben?

Über die Gründe, warum in den Aquificaceae die RNA-basierte RNase P im Laufe der Evolution ersetzt wurde, lässt sich nur spekulieren. Das RNA-basierte Enzym muss koordiniert von zwei Genen synthetisiert und assembliert werden, eine komplexe Biosynthese im Vergleich zur Synthese eines einzelnen Polypeptids. Bei den hohen Wachstumstemperaturen von *Aquifex* (bis 95 °C) sind Probleme bei der RNA-Faltung, der Holoenzym-Assemblierung sowie der Substratbindung denkbar. Allerdings kommen andere thermophile Bakterien wie *Thermoto-*

*ga maritima* (Wachstum bis 90 °C) sehr gut mit einer klassischen RNA-basierten RNase P zurecht. Vielleicht ist der Austausch der klassischen RNase P durch HARP in den Aquificaceae auch nur ein evolutionärer Zufall oder dem evolutionären Druck zuzuschreiben, das *Aquifex*-Genom auf eine möglichst geringe Größe zu kondensieren.

#### Literatur

- [1] Holzmann J, Frank P, Löffler E et al. (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* 135:462–474
- [2] Nickel AI, Wäber NB, Gößringer M et al. (2017) Minimal and RNA-free RNase P in *Aquifex aeolicus*. *Proc Natl Acad Sci USA* 114:11121–11126
- [3] Deckert G, Warren PV, Gaasterland T et al. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* 392:353–358
- [4] Willkomm DK, Feltens R, Hartmann RK (2002) tRNA maturation in *Aquifex aeolicus*. *Biochimie* 84:713–722

- [5] Marszałkowski M, Willkomm DK, Hartmann RK (2008) 5'-end maturation of tRNA in *Aquifex aeolicus*. *Biol Chem* 389:395–403
- [6] Lombo TB, Kaberdin VR (2008) RNA processing in *Aquifex aeolicus* involves RNase E/G and an RNase P-like activity. *Biochem Biophys Res Commun* 366:457–463
- [7] Matelska D, Steczkiewicz K, Ginalski K (2017) Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Res* 45:6995–7020
- [8] Veith T, Martin R, Wurm JP et al. (2012) Structural and functional analysis of the archaeal endonuclease Nob1. *Nucleic Acids Res* 40:3259–3274

#### Korrespondenzadresse:

Prof. Dr. Roland K. Hartmann  
Institut für Pharmazeutische Chemie  
Philipps-Universität Marburg  
Marbacher Weg 6  
D-35037 Marburg  
Tel.: 06421-28-25827  
Fax: 06421-28-25854  
roland.hartmann@staff.uni-marburg.de  
<http://biochemie.pharmazie.uni-marburg.de>

#### AUTOREN



Isabell Schencking, Markus Gößringer, Astrid I. Nickel (Foto im Foto), Roland K. Hartmann und Nadine B. Wäber (v. l. n. r.)

#### Astrid I. Nickel

2002–2006 Pharmaziestudium an der Universität Marburg. 2008 Approbation als Apothekerin. 2008–2014 Promotion an der Universität Marburg. 2014–2015 Postdoc an der Universität Marburg. Seit 2016 Mitarbeiterin der Haupt Pharma Wülfing GmbH (Aenova Group), Gronau/Leine.

#### Isabell Schencking

2011–2015 Pharmaziestudium an der Universität Marburg. 2016 Approbation als Apothekerin. Seit 2017 Promotion am Institut für Pharmazeutische Chemie der Universität Marburg.

#### Nadine B. Wäber

2007–2010 Biologiestudium an der Universität Marburg mit Schwerpunkt Mikrobiologie. 2011–2013 Masterstudium am Karlsruher Institut für Technologie (KIT) mit den Schwerpunkten Genetik, Molekular- und Mikrobiologie. Seit 2014 Promotion am Institut für Pharmazeutische Chemie der Universität Marburg.

#### Markus Gößringer

1990–1996 Biologiestudium (Diplom) an der Universität Konstanz. 1999–2004 Promotion an den Universitäten Lübeck und Marburg. Seit 2004 Postdoc am Institut für Pharmazeutische Chemie der Universität Marburg.

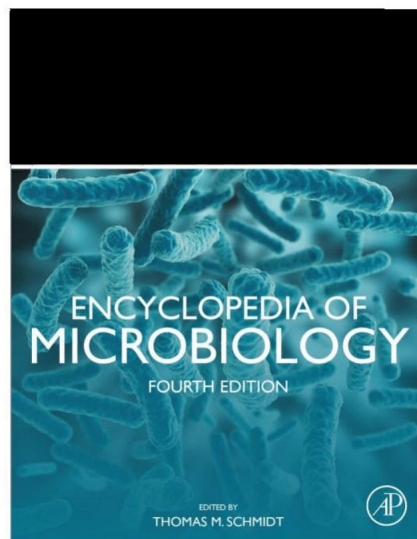
#### Roland K. Hartmann

1976–1981 Biochemiestudium an der FU Berlin, dort 1983–1987 Promotion. 1988 Ernst-Reuter-Preis für hervorragende Dissertationen. 1995–2003 Arbeitsgruppenleiter am Institut für Biochemie der Universität Lübeck. Seit 2003 Professor für Pharmazeutische Chemie an der Universität Marburg.

## Artikel in „Encyclopedia of Microbiology” – 2019

Provided for non-commercial research and educational use.  
Not for reproduction, distribution or commercial use.

This article was originally published in the *Encyclopedia of Microbiology* published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use, including without limitation, use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation, commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<https://www.elsevier.com/authors/about/policies/copyright/permissions>

Wäber Nadine B., and Hartmann Roland K. (2019) Aquificae. In: Schmidt, Thomas M. (ed.) *Encyclopedia of Microbiology*, 4th Edition. vol. 1, pp. 226-233. UK: Elsevier.

[dx.doi.org/10.1016/B978-0-12-809633-8.20783-8](https://doi.org/10.1016/B978-0-12-809633-8.20783-8)

© 2019 Elsevier Inc. All rights reserved.

## Aquificae

Nadine B Wäber and Roland K Hartmann, Philipps-University Marburg, Marburg, Germany

© 2019 Elsevier Inc. All rights reserved.

### Glossary

**crcB** The *crcB* RNA motif is now called the fluoride riboswitch. A *crcB* motif from *Pseudomonas syringae* was shown to be capable of selectively sensing the very small and highly charged fluoride ion.

**Chemolithoautotroph** Organism that uses anorganic compounds as electron donors and carbon source.

**Chemolithoheterotroph** Organism that uses anorganic compounds as electron donors, but organic compounds as carbon source.

**Chemoorganoheterotroph** Organism that uses organic compounds as electron donors and carbon source.

### Abbreviations

Co	Cobalt
Fe	Iron
HGT	Horizontal gene transfer
kbp	Kilo (10 <sup>3</sup> ) base pairs
Mn	Manganese
Moco	Molybdenum cofactor
ORF	Open reading frame
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
SRP RNA	Signal recognition particle RNA
tmRNA	Transfer-messenger RNA

### Introduction

The bacterial phylum Aquificae comprises the orders *Aquificales* (families *Aquificaceae* and *Hydrogenothermaceae*) and *Desulfurobacteriales* (family *Desulfurobacteriaceae*) (Gupta and Lali, 2013). Members of the Aquificae inhabit terrestrial and marine hydrothermal systems around the planet. For example, genera from all three families can be isolated at deep-sea vents such as the Eastern Lau Spreading Center and Valu-Fa Ridge (ELSC/VFR) in the Southwestern Pacific (Ferrera et al., 2014), where they inhabit hydrothermal, mineral and sulfur-rich deposits that form when the cold oxygenated seawater mixes with the high temperature hydrothermal fluid. Such geochemical habitats are dominated by mesophilic  $\epsilon$ -proteobacteria, but Aquificae complement these bacterial communities, evidently by thriving in the higher temperature areas (Ferrera et al., 2014). The phylum members *Aquifex aeolicus* and *Aquifex pyrophilus* are the most temperature-resistant bacteria known; they grow up to 95°C, whereas other *Aquificales* have upper growth temperature limits of ~ 80°C. The known Aquificae are Gram-negative, non-sporulating and motile bacteria that grow at slightly acidic (~ pH 5) to neutral pH and which are adapted to the salinity of their habitats (Guiral et al., 2012). They are primarily chemolithoautotrophs that utilize anorganic compounds as energy source (electron donor) and obtain carbon by CO<sub>2</sub> fixation using the reductive tricarboxylic acid (rTCA) cycle. Electron donors are H<sub>2</sub> (with the exception of some *Hydrogenothermaceae* species), elemental sulfur (S<sup>0</sup>) or thiosulfate (S<sub>2</sub>O<sub>3</sub><sup>2-</sup>); electron acceptors are O<sub>2</sub> in the case of the microaerobic *Aquificaceae* and *Hydrogenothermaceae* (some organisms can also use NO<sub>3</sub><sup>-</sup> or S<sup>0</sup>), whereas the strictly anaerobic *Desulfurobacteriaceae* only utilize sulfur compounds like S<sup>0</sup> or NO<sub>3</sub><sup>-</sup> as electron acceptors. Some characteristics of the eleven Aquificae strains with completely assembled genome sequence (or of related strains if the sequenced strains have not been characterized specifically) plus *A. pyrophilus* as a founding member of the phylum (Huber et al., 1992) are summarized in Table 1.

### Energy Metabolism – Idiosyncrasies, Commonalities and Differences

*Aquifex* stands for “water-maker” to indicate H<sub>2</sub> oxidation and terminal electron transfer to O<sub>2</sub> (“Knallgas” reaction). However, some *Sulfurihydrogenibium* strains such as sp. Y03AOP1, *S. yellowstonense* and sp. UZ 3–5 are unable to oxidize H<sub>2</sub>.



**Table 1** Physiological characteristics of members of the phylum Aquificae

	Growth T (°C) min- opt-max	Electron donor	Electron acceptor	Products of energy metabolism	Carbon source	Other requirements	Cell morphology	Growth style <sup>d</sup>	References
<b>Family Aquificaceae</b>									
<i>Aquifex aeolicus</i> VF5	58-85-95	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup>	O <sub>2</sub> (low), S <sup>0</sup>	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> , H <sub>2</sub> S	CO <sub>2</sub>	mineral salts/trace elements	Gram <sup>+</sup> , rods, motile	cla, ma	(Decker <i>et al.</i> , 1998; Guiral <i>et al.</i> , 2005; Huber and Eder, 2006) (Huber <i>et al.</i> , 1992)
<i>Aquifex pyrophilus</i> Kuls1	67-85-95	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup>	O <sub>2</sub> (low), NO <sub>3</sub> <sup>-</sup>	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> , H <sub>2</sub> S, NO <sub>2</sub> <sup>-</sup> —N <sub>2</sub>	CO <sub>2</sub>			cla, ma, aa	
<sup>b</sup> <i>Hydrogenivirga calditioris</i> IBSK3	55-75-77.5	H <sub>2</sub> , S <sup>0</sup>	O <sub>2</sub> (low), NO <sub>3</sub> <sup>-</sup>	H <sub>2</sub> O, N <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup>	CO <sub>2</sub>			cla, ma, aa	(Nakagawa <i>et al.</i> , 2004; Reysenbach <i>et al.</i> , 2009)
<sup>c</sup> <i>Hydrogenobacter thermophilus</i> TK-6	70 50-71-78 75	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , HCO <sub>3</sub> <sup>-</sup> , HCONH <sub>2</sub>	O <sub>2</sub> (low), NO <sub>3</sub> <sup>-</sup>	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> , H <sub>2</sub> S, NO <sub>2</sub> <sup>-</sup> , NO, N <sub>2</sub> O—N <sub>2</sub>	CO <sub>2</sub>			cla, coh, ma, aa	(Kawasumi <i>et al.</i> , 1984; Suzuki <i>et al.</i> , 2001; Eder and Huber, 2002)
<sup>d</sup> <i>Hydrogenobaculum acidophilum</i>	~45-65-70	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup>	O <sub>2</sub> (low)	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup>	CO <sub>2</sub>			cla, ma	(Shima and Suzuki, 1993; Stöhr <i>et al.</i> , 2001; Reysenbach <i>et al.</i> , 2009)
<sup>e</sup> <i>Thermocrinis ruber</i> DSM 23557	44-80-89	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup> , HCO <sub>3</sub> <sup>-</sup> , HCONH <sub>2</sub> , AsO <sub>4</sub> <sup>3-</sup> , AsSO <sub>4</sub> <sup>3-</sup>	O <sub>2</sub> (low), S <sup>0</sup>	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> , H <sub>2</sub> S, NH <sub>3</sub> , AsO <sub>4</sub> <sup>3-</sup>	CO <sub>2</sub>			cla, coh, ma	(Huber <i>et al.</i> , 1998; Härtig <i>et al.</i> , 2014)
<i>Thermocrinis albus</i> DSM 14484	44-80-89	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup> , HCO <sub>3</sub> <sup>-</sup> , HCONH <sub>2</sub>	O <sub>2</sub> (low), S <sup>0</sup>	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> , H <sub>2</sub> S, NH <sub>3</sub>	CO <sub>2</sub>			cla, coh, ma	(Huber <i>et al.</i> , 1998)
<b>Family Hydrogenothermaceae</b>									
<sup>f</sup> <i>Sulfurhydrogenibium</i> sp. Y03ADP1	~70 (optimal)	S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup>	O <sub>2</sub> (low) <sup>g</sup>	SO <sub>4</sub> <sup>2-</sup>	CO <sub>2</sub>			cla, ch, ma, aa	(Takai <i>et al.</i> , 2003; Nakagawa <i>et al.</i> , 2005; O'Neill <i>et al.</i> , 2008; Reysenbach <i>et al.</i> , 2009)
<sup>h</sup> <i>Sulfurhydrogenibium azorense</i> Az-Fu1	50-68-73	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup> , SO <sub>3</sub> <sup>2-</sup> , Fe <sup>2+</sup> , ASO <sub>3</sub> <sup>3-</sup>	O <sub>2</sub> (low), S <sup>0</sup> , Fe <sup>2+</sup> , SO <sub>3</sub> <sup>2-</sup> , SeO <sub>4</sub> <sup>2-</sup> , AsO <sub>4</sub> <sup>3-</sup>	H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> , Fe <sup>2+</sup> , AsO <sub>4</sub> <sup>3-</sup>	CO <sub>2</sub> yeast extract, bacto peptone, trypticase peptone, casamino acids			cla, ch, ma, aa	(Aguilar <i>et al.</i> , 2004; Nakagawa <i>et al.</i> , 2005)
<i>Persephonella marina</i> EX-H1	55-73-80	H <sub>2</sub> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , S <sup>0</sup>	O <sub>2</sub> (low), NO <sub>3</sub> <sup>-</sup> , S <sup>0</sup>	H <sub>2</sub> O, H <sub>2</sub> S, N <sub>2</sub>	CO <sub>2</sub>			cla, ma	(Götz <i>et al.</i> , 2002)
<b>Family Desulfurobacteriaceae</b>									
<i>Desulfobacterium thermolithotrophum</i> DSM 11699	40-70-75	H <sub>2</sub>	S <sup>0</sup> , S <sub>2</sub> O <sub>3</sub> <sup>2-</sup> , SO <sub>3</sub> <sup>2-</sup>	H <sub>2</sub> S	CO <sub>2</sub>			cla, aa	(L'Haridon <i>et al.</i> , 1998)
<i>Thermovibrio ammonificans</i> HB-1	60-75-80	H <sub>2</sub>	S <sup>0</sup> , NO <sub>3</sub> <sup>-</sup>	H <sub>2</sub> S, NH <sub>4</sub> <sup>+</sup>	CO <sub>2</sub>			cla, aa	(Vetriani <i>et al.</i> , 2004)

<sup>a</sup>cla: chemolithoautotroph (uses inorganic compounds as electron donors and carbon source); coh: chemorganoheterotroph (uses organic compounds as electron donors and carbon source); ch: chemolithoheterotroph (uses inorganic compounds as electron donors, but organic compounds as carbon source); ma: microaerob; aa: anaerobic.

<sup>b</sup>*Hydrogenivirga calditioris* IBSK3 is related to strain *Hydrogenivirga* sp. 128-5-R1-1 (Ferra *et al.*, 2014) that was subjected to genome sequencing (NCBI, NZ\_LBHU000000000.1); another related strain is *Hydrogenivirga okinawensis* LS12-2 (Nunoura *et al.*, 2008) with the following determined features: a growth temperature range of 65-85°C (optimal 70-75°C); use of S<sup>0</sup> and S<sub>2</sub>O<sub>3</sub><sup>2-</sup>, but not H<sub>2</sub> as electron donors; N<sub>2</sub>, S<sub>2</sub>O<sub>3</sub><sup>2-</sup> and SO<sub>4</sub><sup>2-</sup> as products of energy metabolism; strictly chemolithoautotrophic, microaerobic and anaerobic growth (Nunoura *et al.*, 2008).

<sup>c</sup>unable to grow anaerobically with nitrate in the presence of formate, formamide or formaldehyde (Eder and Huber, 2002).

<sup>d</sup>*Hydrogenobaculum acidophilum* (optimal pH: 3-4) is related to *Hydrogenobaculum* sp. Y04AAS1 (optimal pH: 4; Ferra *et al.*, 2007), the strain for which the genome sequence is available. The latter was reported to be unable to reduce nitrate under laboratory conditions (Reysenbach *et al.*, 2009), although genes for nitrate respiration are encoded (Reysenbach *et al.*, 2009; Romano *et al.*, 2013). The optimal growth temperature for strain Y04AAS1 is 58°C (Ferra *et al.*, 2007). A substantial microdiversity of *Hydrogenobaculum* in Yellowstone National Park samples was observed; ~ half of the genes unique to strain Y04AAS1 relative to *Hydrobacterium* sp. isolates reported in (Romano *et al.*, 2013) were predicted to be acquired by HGT. *Hydrobacterium* sp. can further grow on H<sub>2</sub>S as energy source (Romano *et al.*, 2013).

<sup>e</sup>*Thermocrinis ruber* isolate OC14772 (DSM No. 23557) was used as authentic type strain instead of the deposited isolate DSM No. 12173 (Huber *et al.*, 1998) which turned out to be non-authentic based on 16S rRNA sequence analysis (Härtig *et al.*, 2014).

<sup>f</sup>*Sulfurhydrogenibium* (Takai *et al.*, 2003; Aguilar *et al.*, 2004; Nakagawa *et al.*, 2005; O'Neill *et al.*, 2008).

<sup>g</sup>No specific strain characterization available; the given data are based on descriptions of the genus *Sulfurhydrogenibium* (Takai *et al.*, 2003) based on the utilization of the same electron donors/acceptors by *Sulfurhydrogenibium azorense* Az-Fu1 (Aguilar *et al.*, 2004; Nakagawa *et al.*, 2005).

<sup>h</sup>Optimally microaerobic conditions, but aerobic conditions (> 10 volume % O<sub>2</sub>) tolerated by some species.

<sup>i</sup>A novel species, *Desulfurobacterium indicum*, was characterized recently and compared with four other *Desulfurobacterium* species including *D. thermolithotrophum*, which differ in their utilization of terminal electron acceptors (Cao *et al.*, 2017). With H<sub>2</sub> as electron donor, *D. indicum* can utilize S<sup>0</sup>, S<sub>2</sub>O<sub>3</sub><sup>2-</sup> and NO<sub>3</sub><sup>-</sup> as electron acceptors, but not SO<sub>3</sub><sup>2-</sup>.

(O'Neill et al., 2008; Reysenbach et al., 2009) which is consistent with the failure to identify hydrogenase genes in the corresponding genomes. A commonality of all Aquificae is their capability to fix CO<sub>2</sub> by the rTCA cycle to synthesize acetyl-CoA. While the *Aquificaceae* utilize the presumedly ancestral "B-type" rTCA cycle (cleavage of citrate to oxaloacetate and acetyl-CoA by successive action of two enzymes, citryl-CoA synthetase and citryl-CoA lyase), the "A-type" rTCA cycle (cleavage of citrate to oxaloacetate and acetyl-CoA by the single two-subunit enzyme ATP citrate lyase) is implemented in the other two Aquificae families (Hügler et al., 2007). The *Desulfurobacteriaceae* additionally encode an incomplete reductive acetyl-CoA pathway thought to be the most ancient carbon fixation pathway on earth. Together with their ability to utilize sulfur as electron acceptor and their strictly anaerobic growth, this family was proposed to have retained primordial metabolic components that evolved before the emergence of photosynthesis and the rise of oxygen in the atmosphere (Giovannelli et al., 2017). A common trait of members of the order *Aquificales* is their ability to utilize elemental sulfur or inorganic sulfur compounds as electron donors (Table 1).

### Oxidative Stress

Aquificae that grow under microaerobic conditions have to cope with reactive oxygen species (ROS) such as superoxide and peroxide. Three genes encoding superoxide dismutase (one Fe/Mn and two Co/Zn family enzymes) were identified in the *A. aeolicus* genome. However, no catalase gene could be identified, although *Aquifex* cells react with H<sub>2</sub>O<sub>2</sub> as inferred from rapid gas production (Huber et al., 1992; Swanson, 2001). Catalase activity, detected in the insoluble fraction of *A. aeolicus* cell lysates, was resistant to acid treatment and could not be extracted by organic solvents from insoluble lysate pellets. Further analysis identified phosphate salts of Fe as a major component of such insoluble fractions. As Fe alone has catalase activity, it has been speculated that *A. aeolicus* may promote the intracellular precipitation of such inorganic Fe-phosphate salts to protect cells against damage by H<sub>2</sub>O<sub>2</sub> (Swanson, 2001).

*Thermovibrio ammonificans*, which grows strictly anaerobic, also encodes proteins that protect against damage by ROS. This includes a catalase/peroxidase, a putative superoxide reductase, a cytochrome c peroxidase and a cytochrome bd complex. The latter was shown to contribute to oxygen tolerance in anaerobic bacteria and to detoxification of nitrous oxide radicals (Giovannelli et al., 2017).

### Gluconeogenesis

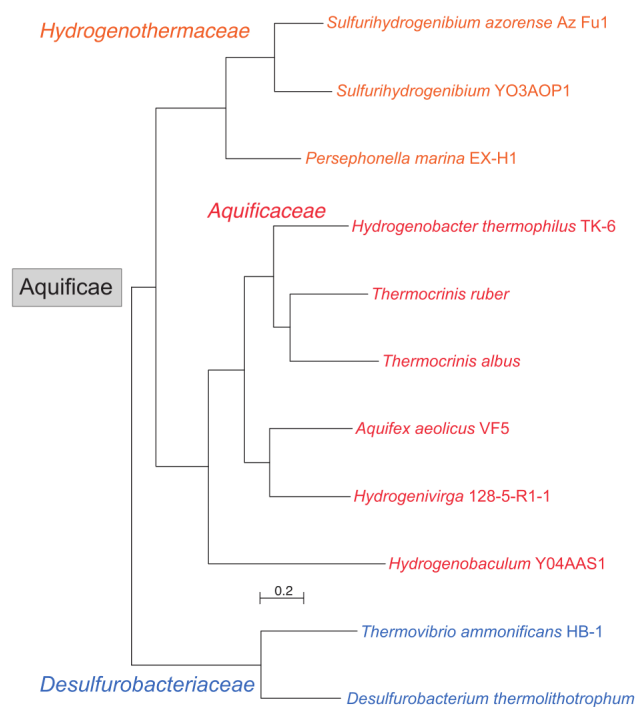
The Aquificae synthesize pentose and hexose monosaccharides from products of the rTCA cycle by the Embden-Meyerhof-Parnas pathway. The key gluconeogenic enzyme fructose-1,6-bisphosphatase (*fbp*) was identified in *T. ammonificans*, *Desulfobacterium thermolithotrophum*, *Sulfurihydrogenibium* sp. Y03AOP1, *Persephonella marina*, *Hydrogenivirga* 128-5-R-1-1, but not in *Hydrogenobacter thermophilus* and *A. aeolicus*, suggesting a pathway variation in the latter two bacteria (Deckert et al., 1998; Giovannelli et al., 2017). Genes encoding enzymes of the pentose-phosphate pathway and for glycogen synthesis and catabolism were also found in the *A. aeolicus* genome (Deckert et al., 1998).

### Motility

All Aquificae are motile, forming one or two (mono- or polytrichous) monopolar flagella. Some members of the phylum, such as *T. ammonificans* and *D. thermolithotrophum*, also encode homologs of the bacterial chemotaxis system (Giovannelli et al., 2017); corresponding genes were not identified in the *A. aeolicus* genome (Swanson, 2001). As the well-characterized bacterial chemotaxis systems respond to the presence of sugars and amino acids, chemolithoautotrophs like *A. aeolicus* may sense other changes, such as concentration of dissolved CO<sub>2</sub>, H<sub>2</sub>, O<sub>2</sub> or temperature, by other mechanisms (Swanson, 2001).

### Sizes and G/C Content of Aquificae Chromosomes

To date (end of 2018), the NCBI genome browser lists 23 Aquificae genome projects. Ten represent complete genome assemblies, all others are in premature assembly states. We also included *Hydrogenivirga* in our comparative genome analyses, as done in a recent study to which we refer here (Lechner et al., 2014; see phylogenetic tree in Fig. 1). The eleven strains are specified in Table 1; if the strain with sequenced genome was not characterized in detail in the literature, we instead describe the features of a closely related strain. In addition, we included the founding strain *A. pyrophilus* with unsequenced genome for comparison of metabolic features (Table 1). Genome sizes are close to the assumed lower limit of genomes in free-living bacteria, ranging from 1.5 (*T. albus*) to ~2 Mbp (*P. marina*). The current annotation file of the *Hydrogenivirga* 128-5-R-1-1 genome contains a sequence of 3.04 Mbp, which is roughly double the size as found in other Aquificae. However, the *Hydrogenivirga* sequence file only represents the lowest contig assembly level and there is evidence that the genome assembly is erroneous, likely a blend of two genomes of related bacteria, one from a member of the *Aquificaceae* and the second from a *Hydrogenothermaceae* species. This is inferred from the finding of two 6S



**Fig. 1** Maximum likelihood tree of Aquificae based on whole genome alignments. Strains: *Thermocrinis ruber* DSM 23557; *Thermocrinis albus* DSM 14484; *Desulfurobacterium thermolithotrophum* DSM 11699. Reproduced from Lechner, M., Nickel, A.I., Wehner, S., 2014. Genomewide comparison and novel ncRNAs of Aquificales. BMC Genomics 15, 522. doi:10.1186/1471-2164-15-522.

RNA candidates and two different subtype tmRNAs – rather than single loci as in the other Aquificae – as well as an increased number of tRNA genes and CRISPR loci (Lechner et al., 2014). The genome G/C content varies between 32% (*Sulfurihydrogenibium* sp. Y03AOP1) to 52.1% (*T. ammonificans*).

### Extrachromosomal Plasmids

Plasmids were isolated and sequenced for *A. aeolicus* (39 kbp, 32 genes), *T. ammonificans* (77 kbp, 97 genes) and *P. marina* (54 kbp, 74 genes). In *A. aeolicus*, on average two plasmid copies were estimated to be present per chromosome. Except for a transposase, ORFs were classified as hypothetical proteins; one plasmid ORF is also present in two identical copies on the chromosome (Deckert et al., 1998). In *T. ammonificans*, the plasmid encodes mostly ORFs with unknown hypothetical functions, except for a putative RNA polymerase sigma factor, a DNA topoisomerase I, a type II secretion protein E and an ArsR-like helix-turn-helix transcription regulator. Furthermore, the plasmid contains highly repeated DNA (Giovannelli et al., 2017). In *P. marina*, apart from mostly hypothetical proteins, the plasmid encodes, among others, an antitoxin, several proteins involved in conjugal plasmid transfer, a site-specific integrase, a type IA DNA topoisomerase, a plasmid replication initiator and segregation proteins.

### Protein Genes and mRNA Decoding

The first published Aquificae genome of *A. aeolicus* revealed several idiosyncrasies (Deckert et al., 1998). In the densely packed genome, most genes are part of polycistronic transcription units as in other Bacteria. There are many operons encoding genes of the



same functional context, which is pronounced for ribosomal protein genes, but also seen for many genes involved in energy metabolism, cell envelope biogenesis and processes such as flagella synthesis. Yet, there are also many genes of the same functional context, such as those of the tryptophan and histidine biosynthesis pathways, that are not organized in the same operons as in other Bacteria, but are dispersed in the *A. aeolicus* genome or appear in novel operons. This also pertains to genes for subunits of the same enzyme, for example *gluB* (Aq\_1565) and *gluD* (Aq\_2064; glutamate synthase), the  $\alpha$  and  $\beta$  subunit genes of glycyl and phenylalanine tRNA synthetases (Aq\_945/2141 and Aq\_953/1730), *ribH/ribC* (Aq\_132/1707) encoding the  $\alpha$  and  $\beta$  subunits of riboflavin synthase and *nrdA/nrdF* (Aq\_94/1505) coding for the  $\alpha$  and  $\beta$  chains of ribonucleotide reductase. It has been surmised that the lack of cistronic arrangement of biosynthetic genes may reflect the chemolithoautotrophic life style of Aquificae that is committed to CO<sub>2</sub> fixation. In contrast, chemoorganoheterotrophic microorganisms can metabolically switch between different energy and carbon sources. This is achieved by coordinated up- and down-regulation of transcription units that encode the genes participating in the respective metabolic pathways (Swanson, 2001).

It is also not clear if transcription in the Aquificae may differ from what is known from other bacteria. The RNA polymerase and sigma factors in the genus *Aquifex* have functional properties very similar to those of other bacteria (Klenk et al., 1999; Studholme and Buck, 2000; Studholme et al., 2000), but *A. aeolicus* was reported to lack hairpin structures for transcription termination (Washio et al., 1998). The latter observation might suggest that transcript ends/operon boundaries are less defined, possibly pointing to read-through transcription of longer genomic segments. However, this may not pertain to the entire genome, as potential (though not very stable) hairpin structures were identified at the end of the consecutive *rplA/L* (encoding ribosomal proteins) and *rpoBC* (encoding the  $\beta$  and  $\beta'$  subunits of RNA polymerase) operons in *A. pyrophilus* (Klenk et al., 1999). Many convergently transcribed ORFs in *A. aeolicus* overlap, ranging from simple stop codon overlaps to overlaps of tens of base pairs (Swanson, 2001). In RNA-Seq experiments, high sense and antisense read levels were found in genes encoding validated and putative non-coding RNAs. The same pertains to overlapping 5'-upstream regions of two protein-coding genes located on opposing strands, as well as areas where two 3'-untranslated regions overlap on opposing strands (Lechner et al., 2014). Such RNA reads might be vestiges of extensive transcription of both genome strands, although this remains to be validated.

Whole genome alignments using the software tools Pomago (v. 1.0) and TBA (v. 11.2) revealed that individual genomes have gaps of 10.5%–13% (Pomago) or 8.4%–9.6% (TBA) when aligned to the other ten genomes. The original protein annotations of the genome files were recently extended, predicting for example  $\sim 6\%$  more proteins with assigned or hypothetical function in *A. aeolicus*. In addition, the 5'- and 3'-ends of ORFs were partially refined. The Shine-Dalgarno sequences in the 11 organisms are quite conserved (5'-GGAGG, always 5'-NGAGN), -10 promoter signatures are more variable though recognizable (A/T-rich). The number of tRNA genes is minimized to 39–46 (*Hydrogenivirga* excluded) compared with  $> 80$  tRNA genes in strains of *Escherichia coli* and *Bacillus subtilis*. With respect to codon usage, the two Arg codons with 5'-A and the Ile AUA codon are preferentially utilized in the Aquificaceae and Hydrogenothermaceae. The preference for AUA correlates with expression of tRNA-Ile lysidine synthetase (TilS) in the Aquificae (a crystal structure of *A. aeolicus* TilS is available). TilS converts the 5'-C residue in the CAU anticodon of specific tRNA-Ile molecules to lysidine (2-lysyl cytidine; abbreviated as L or  $k^2C$ ) to allow decoding of 5'-AUA (Ile) instead of 5'-AUG (Met) codons. Apparently, this anticodon modification enables reliable decoding at high temperatures. Consistent with the reduced number of tRNA genes (39–46) in the Aquificae, multiple copies of tRNA isoacceptor genes are avoided. Almost no tRNAs with 5'-A in the anticodon are encoded and anticodons with 5'-C are reduced in the non-Aquificaceae within the Aquificae. The streamlining of the tRNA repertoire goes along with expression of the tRNA modification enzyme tRNA adenosine deaminase (TadA; crystal structure of the *A. aeolicus* protein available) that converts the anticodon 5'-A residue to inosine in the single tRNA<sup>Arg</sup> (anticodon 5'-ACG). This allows decoding of codons 5'-CG(U,C,A), while each of the three other tRNA<sup>Arg</sup> isoacceptors recognize one of the remaining three Arg codons through Watson-Crick base-pairing at all three codon positions. The maintenance of wobble decoding by tRNA<sup>Arg</sup> (5'-ICG) while decreasing the fraction of 5'-CG(U,C,A) codons in the genome may be a compromise resulting from constraints to downsize the tRNA repertoire and to minimize the number of labile codon:anticodon interactions. The Aquificaceae (except for *Hydrogenobaculum*) and *Desulfurobacteriaceae* encode selenocysteine-specific tRNAs decoding 5'-UGA codons; those are absent from the *Hydrogenothermaceae*. The Aquificaceae, but not *Hydrogenivirga* and the other Aquificae, express the lysine tRNA isoacceptor with the anticodon 5'-CUU to stabilize AAG decoding, a feature shared with other thermophilic bacteria (*Thermus thermophilus*, *Thermotoga maritima*). In mesophiles, AAG codons are decoded by isoacceptors harboring the anticodon 5'-UUU.

### Non-Coding RNAs (ncRNAs) in the Aquificae

Apart from tRNAs and 1–3 copies of 5S/16S/23S rRNA, bioinformatics identified genes for 6S RNA (except for uncertain candidates in *D. thermolithotrophum* and *T. ammonificans*), a regulator of bacterial RNA polymerase, the signal recognition particle RNA (SRP RNA), and a tmRNA of type A or B; riboswitch candidates were found sporadically in a few species [TPP, Moco, Cobalamin and *crcB* (Lechner et al., 2014)]. Of note, 6S RNA is present in the Aquificaceae and Hydrogenothermaceae, as well as in *Thermodesulfatator indicus* of the Aquificae sister phylum Thermodesulfobacteria, but is absent from the next related phyla (according to 16S rRNA phylogeny), Thermotogae and Deinococcus-Thermus. RNase P RNA is not encoded in the genomes of Aquificaceae, but is present in the two other families. Instead, the Aquificaceae encode a novel type of protein-only RNase P, a PIN\_5 domain metallo-nuclease consisting of a single 23-kDa polypeptide. The genes for the ancient RNA-based enzyme (*mpB* encoding the RNA and *mpA* the protein subunit), present in the vast majority of Bacteria, were lost in the Aquificaceae, which represents a novel hallmark feature of

this family. Homologs of *Aquifex* RNase P (HARP) were identified in many Archaea and in a few other bacteria of different phyla; all these Archaea encode an RNA-based RNase P as well. Combined with the assessment that at least 10% of genes in *A. aeolicus* were acquired by horizontal gene transfer (HGT) from Archaea (Aravind et al., 1998), it is very likely that the progenitor of the *Aquificaceae* acquired the HARP gene by HGT from an archaeon (Nickel et al., 2017). In accordance, bootstrap analysis groups all bacterial HARPs with those of archaeal Thermococci, suggesting that bacterial HARPs originate from this class of Euryarchaeota (Nickel et al., 2017).

The *A. aeolicus* genome was analyzed for ncRNA candidates, applying constraints such as a minimum length of 25 nt and no overlap with protein-coding sequences, rRNA or tRNA genes. Since the housekeeping ncRNAs discussed above have a G/C content of 66% in *A. aeolicus*, an elevated G/C content (relative to the 43% average G/C content of the genome) was taken as an additional criterion for ncRNA identity. The analysis identified a candidate list of ~ 100 ncRNAs in *A. aeolicus* (Lechner et al., 2014), including the housekeeping ncRNAs, such as SRP RNA, tmRNA or 6S RNA. As mentioned before, RNA-Seq data for *A. aeolicus* suggest a substantial level of antisense transcripts for known ncRNAs and novel ncRNA candidates, which could be partly confirmed by Northern blot analyses (Lechner et al., 2014).

### Mechanisms of Thermostabilization

Adaptations to growth at high temperatures are evident at the level of cell envelope, genome, RNAs and proteins. The cell membranes of several *Aquificaceae* were found to contain the more stable glycerol-ether phospholipids in addition to acyl glycerides (Jahnke et al., 2001). *A. aeolicus* expresses two reverse gyrases (genes Aq\_886 and Aq\_1159), which are topoisomerases originally acquired from Archaea by HGT and which introduce positive supercoils into the DNA. Potential functions include prevention of excess local unwinding of the double helix or protection of DNA against damage (e.g., depurination, strand breakage) at high temperature (Brochier-Armanet and Forterre, 2007; Ogawa et al., 2015). Another adaptation to (hyper)thermophilic conditions is an elevated G/C content of non-coding RNAs (ncRNAs). The ribosomal RNAs of *A. aeolicus* have 65% G/C compared with an average G/C content of ~ 43% in the bacterium's genomic DNA (Deckert et al., 1998). This G/C-richness, which correlates with enhanced RNA secondary and tertiary structural stability, is also implemented in other ncRNAs of *A. aeolicus*, such as tRNAs, tmRNA or 6S RNA, the thermodynamically most stable 6S RNA known so far (Deckert et al., 1998; Willkomm et al., 2005; Köhler et al., 2015). Increased thermostability of proteins is achieved by multiple adaptations that are not specific for the Aquificae. Of note, alterations relative to mesophilic Bacteria and Archaea differ between moderately and extremely thermophilic organisms, and different protein families have evolved individual strategies to achieve high thermal stability (Szilágyi and Závodszky, 2000). Hyperthermophiles (Archaea and Bacteria such as *A. aeolicus*) have increased proportions of charged amino acids (Lys, Glu) at the expense of polar non-charged amino acids (Asn, Gln, Ser, Thr), indicative of an increase in ion pairs including salt bridges between N- and C-termini (Cambillau and Claverie, 2000; Szilágyi and Závodszky, 2000). Disulfide bonds can also contribute to stabilization, such as in an *Aquifex* serine protease containing eight Cys residues (Choi et al., 1999), ferredoxins (Meyer et al., 2002) or rhodanese (Aq\_1599, (Giuliani et al., 2010)). There is a trend to increase the fraction of  $\alpha$ -helices and  $\beta$ -strands and to decrease flexible regions (e.g., surface loops) in thermophilic proteins. An example is a thermostable cytochrome c equipped with an extra  $\alpha$ -helix (Obuchi et al., 2009). C-terminal ends of proteins are frequently folded into the protein core to avoid flexible and accessible regions.

Apart from increases in the number of ion pairs, increases in the number of hydrogen bonds or enhanced core hydrophobicity have been questioned as substantial contributors to extreme thermostability (Szilágyi and Závodszky, 2000). Yet, protein oligomerization is often enhanced. In an archaeal histone dimer, improved intermolecular hydrophobic interactions and additional favorable ion pairs at or near the dimer interface were found to be associated with thermostability, along with the preference for large hydrophobic side chains to reduce cavity sizes and to increase packaging density in the hydrophobic dimer core; in addition, C-terminal protein extension by two Lys residues was proposed to shield the dimer core from solvent exposure (Li et al., 2000). As another example, an *A. aeolicus* rhodanese enzyme (Aq\_477) involved in sulfur transfer forms tetramers as a prerequisite for thermostability and enzymatic function. Formation of supercomplexes is another strategy to enhance thermostability. The thermostable sulfide-oxidase and oxygen-reductase supercomplex of *A. aeolicus* was found to be resistant to denaturation/inactivation by up to 8 M urea, 6% Triton X-100 and 1% SDS (Guiral et al., 2012).

### Phylogeny

The phylum Aquificae has been defined by 16S rRNA phylogeny, yet its member organisms show phylogenetic, ecological, morphological and metabolic diversity. Several protein-based phylogenies have supported 16S RNA-based phylogenies that place the Aquificae as a deep branching sister phylum of the Thermotogae in the bacterial phylogenetic tree. Other analyses have indicated a closer affiliation with other bacterial groups such as the  $\epsilon$ -proteobacteria. This is further complicated by bioinformatic evidence that *A. aeolicus* and other Aquificae as well as Thermotogae have acquired many of their protein-coding genes by HGT from Archaea (Aravind et al., 1998; Zhaxybayeva et al., 2009). Two major hypotheses have been formulated to explain the findings of various studies: (a) the Aquificae and the Thermotogae represent the deepest-branching bacterial phyla containing many hyperthermophiles,

with strong affiliations to the  $\epsilon$ -proteobacteria owing to large-scale gene sharing, and further shaped by acquisition of genes from archaeal thermophiles; (b) the Aquificae are essentially  $\epsilon$ -proteobacteria or a sister to this group that extensively exchanged genes with other thermophilic lineages such as Thermotogae and Archaea. The large extent of genetic mosaicism, also observed in other Aquificae such as *T. ammonificans* (Giovannelli et al., 2017), led Eveleigh et al. (2013) to conclude that "The evolutionary history of the Aquificae implicates different partner lineages, most notably the Archaea, Thermotogae,  $\delta$ -proteobacteria, and thermophilic members of Nitrospirae, Clostridia and  $\epsilon$ -proteobacteria, with different lineages making disproportionate contributions to different molecular subsystems." Indeed, bioinformatic analysis of *A. aeolicus* proteins from selected functional subsystems (cell wall, flagella, oxidative phosphorylation, ribosomal proteins) revealed different "closest neighboring phyla or classes", which even differed between the various complexes participating in oxidative phosphorylation (Eveleigh et al., 2013). This mosaic-like origin of genes in a single metabolic pathway also pertains to enzymes involved in the putative reductive acetyl-CoA pathway in *T. ammonificans* (Giovannelli et al., 2017). Eveleigh et al. (2013) favored a scenario according to which most genes with closest relatedness to homologs in thermophiles are derived from HGT, whereas genes related to those in  $\epsilon$ -proteobacteria are relics of a mesophilic past that preceded colonization of high temperature environments. This takes evidence into account that  $\epsilon$ -proteobacteria developed diverse strategies to colonize many deep-sea hydrothermal habitats, favored by their high growth rates, rapid adaptations to changing geochemical conditions and metabolic versatility. Finally, it cannot be excluded that the extensive genetic mosaicism in the Aquificae might have erased the traces of their real ancestry.

## References

- Aguiar P, Beveridge TJ, and Reysenbach A-L (2004) *Sulfurihydrogenibium azorense*, sp. nov., a thermophilic hydrogen-oxidizing microaerophile from terrestrial hot springs in the Azores. *Int. J. Syst. Evol. Microbiol.* 54: 33–39. <https://doi.org/10.1099/ijs.0.02790-0>.
- Arvind L, Tatusov RL, Wolf YI, Walker DR, and Koonin EV (1998) Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. *Trends Genet.* 14: 442–444.
- Brochier-Armanet C and Forterre P (2007) Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggests a complex history of vertical inheritance and lateral gene transfers. *Archaea* 2: 83–93.
- Cambillau C and Claverie JM (2000) Structural and genomic correlates of hyperthermostability. *J. Biol. Chem.* 275: 32383–32386.
- Cao J, Birien T, Gayet N, et al. (2017) *Desulfurobacterium indicum* sp. nov., a thermophilic sulfur-reducing bacterium from the Indian Ocean. *Int. J. Syst. Evol. Microbiol.* 67: 1665–1668. <https://doi.org/10.1099/ijs.0.001837>.
- Choi IG, Bang WG, Kim SH, and Yu YG (1999) Extremely thermostable serine-type protease from *Aquifex pyrophilus*. Molecular cloning, expression, and characterization. *J. Biol. Chem.* 274: 881–888.
- Deckert G, Warren PV, Gaasterland T, et al. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* 392: 353–358.
- Eder W and Huber R (2002) New isolates and physiological properties of the Aquificales and description of *Thermocrinis albus* sp. nov. *Extremophiles* 6: 309–318. <https://doi.org/10.1007/s00792-001-0259-y>.
- Eveleigh RJ, Meehan CJ, Archibald JM, and Beiko RG (2013) Being *Aquifex aeolicus*: Untangling a hyperthermophile's checkered past. *Genome Biol. Evol.* 5: 2478–2497. <https://doi.org/10.1093/gbe/evt195>.
- Ferrera I, Longhorn S, Banta AB, et al. (2007) Diversity of 16S rRNA gene, ITS region and *acdB* gene of the Aquificales. *Extremophiles* 11: 57–64. <https://doi.org/10.1007/s00792-006-0009-2>.
- Ferrera I, Banta AB, and Reysenbach A-L (2014) Spatial patterns of Aquificales in deep-sea vents along the Eastern Lau Spreading Center (SW Pacific). *Syst. Appl. Microbiol.* 37: 442–448. <https://doi.org/10.1016/j.syapm.2014.04.002>.
- Giovannelli D, Sievert SM, Hügler M, et al. (2017) Insight into the evolution of microbial metabolism from the deep-branching bacterium, *Thermovibrio ammonificans*. *Elife* 6: e18990. <https://doi.org/10.7554/elife.18990>.
- Giuliani MC, Jourlin-Castelli C, Leroy G, Hachani A, and Giudici-Ortoni MT (2010) Characterization of a new periplasmic single-domain rhodanese encoded by a sulfur-regulated gene in a hyperthermophilic bacterium *Aquifex aeolicus*. *Biochimie* 92: 388–397. <https://doi.org/10.1016/j.biochi.2009.12.013>.
- Götz D, Banta A, Beveridge TJ, et al. (2002) *Persephonella marina* gen. nov., sp. nov. and *Persephonella guaymasensis* sp. nov., two novel, thermophilic, hydrogen-oxidizing microaerophiles from deep-sea hydrothermal vents. *Int. J. Syst. Evol. Microbiol.* 52: 1349–1359. <https://doi.org/10.1099/ijs.0.02126-0>.
- Guiral M, Tron P, Aubert C, et al. (2005) A membrane-bound multienzyme, hydrogen-oxidizing, and sulfur-reducing complex from the hyperthermophilic bacterium *Aquifex aeolicus*. *J. Biol. Chem.* 280: 42004–42015. <https://doi.org/10.1074/jbc.M508034200>.
- Guiral M, Prunetti L, Aussignargues C, et al. (2012) The hyperthermophilic bacterium *Aquifex aeolicus*: from respiratory pathways to extremely resistant enzymes and biotechnological applications. *Adv. Microb. Physiol.* 61: 125–194. <https://doi.org/10.1016/B978-0-12-394423-8.00004-4>.
- Gupta RS and Lal R (2013) Molecular signatures for the phylum Aquificae and its different clades: proposal for division of the phylum Aquificae into the emended order Aquificales, containing the families Aquificaceae and Hydrogenothermaceae, and a new order Desulfurobacteriales ord. nov., containing the family Desulfurobacteriaceae. *Antonie Van Leeuwenhoek* 104: 349–368. <https://doi.org/10.1007/s10482-013-9957-6>.
- Härtig C, Lohmayer R, Kolt S, et al. (2014) Chemolithotrophic growth of the aerobic hyperthermophilic bacterium *Thermocrinis ruber* OC 14/72 on monotheiosenate and arsenite. *FEMS Microbiol. Ecol.* 90: 747–760.
- Huber R and Eder W (2006) Aquificales. In: M. Dworkin M, Falcow S, Rosenberg E, Schleifer KH, and Stackebrandt E (eds.). *The Prokaryotes. Proteobacteria: Delta, epsilon subclass*, 7: pp. 925–938. New-York: Springer-Verlag.
- Huber R, Wilham T, Huber D, et al. (1992) *Aquifex pyrophilus* gen. nov., sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. *Syst. Appl. Microbiol.* 15: 340–351.
- Huber R, Eder W, Heldwein S, et al. (1998) *Thermocrinis ruber* gen. nov., sp. nov., a pink-filament-forming hyperthermophilic bacterium isolated from Yellowstone National Park. *Appl. Environ. Microbiol.* 64: 3576–3583.
- Hügler M, Huber H, Molyneux SJ, Vetriani C, and Sievert SM (2007) Autotrophic CO<sub>2</sub> fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum Aquificae: Evidence for two ways of citrate cleavage. *Environ. Microbiol.* 9: 81–92.
- Jahnke LL, Eder W, Huber R, et al. (2001) Signature lipids and stable carbon isotope analyses of Octopus Spring hyperthermophilic communities compared with those of Aquificales representatives. *Appl. Environ. Microbiol.* 67: 5179–5189.
- Kawasumi T, Igarashi Y, Kodama T, and Minoda Y (1984) *Hydrogenobacter thermophilus* gen. nov., sp. nov., an extremely thermophilic, aerobic, hydrogen-oxidizing bacterium. *Int. J. Syst. Bacteriol.* 34: 5–10.

- Klerk HP, Meier TD, Durovic P, et al. (1999) RNA polymerase of *Aquifex pyrophilus*: implications for the evolution of the bacterial *rhoBC* operon and extremely thermophilic bacteria. *J. Mol. Evol.* 48: 528–541.
- Köhler K, Duchardt-Ferner E, Lechner M, et al. (2015) Structural and mechanistic characterization of 6S RNA from the hyperthermophilic bacterium *Aquifex aeolicus*. *Biochimie* 117: 72–86. <https://doi.org/10.1016/j.biochi.2015.03.004>.
- L'Hardon S, Cilia V, Messner P, et al. (1998) *Desulfurobacterium thermolithotrophum* gen. nov., sp. nov., a novel autotrophic, sulphur-reducing bacterium isolated from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* 48: 701–711.
- Lechner M, Nickel AI, and Wehner S (2014) Genomewide comparison and novel ncRNAs of *Aquificales*. *BMC Genomics* 15: 522. <https://doi.org/10.1186/1471-2164-15-522>.
- Li WT, Shriver JW, and Reeve JN (2000) Mutational analysis of differences in thermostability between histones from mesophilic and hyperthermophilic archaea. *J. Bacteriol.* 182: 812–817.
- Meyer J, Clay MD, Johnson MK, et al. (2002) A hyperthermophilic plant-type [2Fe-2S] ferredoxin from *Aquifex aeolicus* is stabilized by a disulfide bond. *Biochemistry* 41: 3096–3108.
- Nakagawa S, Nakamura S, Inagaki F, et al. (2004) *Hydrogenivirga calditioris* gen. nov., sp. nov., a novel extremely thermophilic, hydrogen- and sulfur-oxidizing bacterium from a coastal hydrothermal field. *Int. J. Syst. Evol. Microbiol.* 54: 2079–2084. <https://doi.org/10.1099/ijs.0.03031-0>.
- Nakagawa S, Shtaih Z, Banta A, et al. (2005) *Sulfurihydrogenibium yellowstonense* sp. nov., an extremely thermophilic, facultatively heterotrophic, sulfur-oxidizing bacterium from Yellowstone National Park, and emended descriptions of the genus *Sulfurihydrogenibium*, *Sulfurihydrogenibium subterraneum* and *Sulfurihydrogenibium azorense*. *Int. J. Syst. Evol. Microbiol.* 55: 2263–2268. <https://doi.org/10.1099/ijs.0.63708-0>.
- Nickel AI, Wäber NB, Gößringer M, et al. (2017) Minimal and RNA-free RNase P in *Aquifex aeolicus*. *Proc. Natl. Acad. Sci. USA* 114: 11121–11126. <https://doi.org/10.1073/pnas.1707862114>.
- Nunoura T, Miyazaki M, Suzuki Y, Takai K, and Horikoshi K (2008) *Hydrogenivirga okinawensis* sp. nov., a thermophilic sulfur-oxidizing chemolithoautotroph isolated from a deep-sea hydrothermal field, Southern Okinawa Trough. *Int. J. Syst. Evol. Microbiol.* 58: 676–681. <https://doi.org/10.1099/ijs.0.64615-0>.
- Obuchi M, Kawahara K, Motooka D, et al. (2009) Hyperstability and crystal structure of cytochrome c(555) from hyperthermophilic *Aquifex aeolicus*. *Acta Crystallogr. Sect. D Struct. Biol.* 65(Pt 8): 804–813. <https://doi.org/10.1107/S0907444909017314>.
- Ogawa T, Yogo K, Furuke S, et al. (2015) Direct observation of DNA overwinding by reverse gyrase. *Proc. Natl. Acad. Sci. USA* 112: 7495–7500. <https://doi.org/10.1073/pnas.1422031112>.
- O'Neill AH, Liu Y, Ferrera I, Beveridge TJ, and Reysenbach A-L (2008) *Sulfurihydrogenibium rodmanii* sp. nov., a sulfur-oxidizing chemolithoautotroph from the Uzon Caldera, Kamchatka Peninsula, Russia, and emended description of the genus *Sulfurihydrogenibium*. *Int. J. Syst. Evol. Microbiol.* 58: 1147–1152. <https://doi.org/10.1099/ijs.0.65431-0>.
- Reysenbach A-L, Hamamura N, Podar M, et al. (2009) Complete and draft genome sequences of six members of the *Aquificales*. *J. Bacteriol.* 191: 1992–1993. <https://doi.org/10.1128/JB.01645-08>.
- Romano C, D'Imperio S, Woyke T, et al. (2013) Comparative genomic analysis of phylogenetically closely related *Hydrogenobaculum* sp. isolates from Yellowstone National Park. *Appl. Environ. Microbiol.* 79: 2932–2943. <https://doi.org/10.1128/AEM.03591-12>.
- Shima S and Suzuki KI (1993) *Hydrogenobacter acidophilus* sp. nov., a thermoacidophilic, aerobic, hydrogen-oxidizing bacterium requiring elemental sulfur for growth. *Int. J. Syst. Bacteriol.* 43: 703–708.
- Stöhr R, Waberski A, Völker H, Tindall BJ, and Thomm M (2001) *Hydrogenothermus marinus* gen. nov., sp. nov., a novel thermophilic hydrogen-oxidizing bacterium, recognition of *Calderobacterium hydrogenophilum* as a member of the genus *Hydrogenobacter* and proposal of the reclassification of *Hydrogenobacter acidophilus* as *Hydrogenobaculum acidophilum* gen. nov., comb. nov., in the phylum 'Hydrogenobacter/Aquifex'. *Int. J. Syst. Evol. Microbiol.* 51: 1853–1862.
- Studholme DJ and Buck M (2000) The alternative sigma factor sigma(28) of the extreme thermophile *Aquifex aeolicus* restores motility to an *Escherichia coli* *flaA* mutant. *FEMS Microbiol. Lett.* 191: 103–107.
- Studholme DJ, Wigneshweraraj SR, Gallegos MT, and Buck M (2000) Functionality of purified sigma(N) (sigma(54)) and a NifA-like protein from the hyperthermophile *Aquifex aeolicus*. *J. Bacteriol.* 182: 1616–1623.
- Suzuki M, Cui ZJ, Ishii M, and Igarashi Y (2001) Nitrate respiratory metabolism in an obligately autotrophic hydrogen-oxidizing bacterium, *Hydrogenobacter thermophilus* TK-6. *Arch. Microbiol.* 175: 75–78.
- Swanson RV (2001) Genome of *Aquifex aeolicus*. *Methods Enzymol.* 330: 158–169.
- Szllágyi A and Závodszy P (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: Results of a comprehensive survey. *Structure* 8: 493–504.
- Takai K, Kobayashi H, Nealeon KH, and Horikoshi K (2003) *Sulfurihydrogenibium subterraneum* gen. nov., sp. nov., from a subsurface hot aquifer. *Int. J. Syst. Evol. Microbiol.* 53: 823–827. <https://doi.org/10.1099/ijs.0.02506-0>.
- Vetriani C, Speck MD, Elor SV, Lutz RA, and Starovoytov V (2004) *Thermovibrio ammonificans* sp. nov., a thermophilic, chemolithotrophic, nitrate-ammonifying bacterium from deep-sea hydrothermal vents. *Int. J. Syst. Evol. Microbiol.* 54: 175–181. <https://doi.org/10.1099/ijs.0.02781-0>.
- Washio T, Sasayama J, and Tomita M (1998) Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination. *Nucleic Acids Res.* 26: 5456–5463.
- Willkomm DK, Minnerup J, Hüttenhofer A, and Hartmann RK (2005) Experimental RNomics in *Aquifex aeolicus*: identification of small non-coding RNAs and the putative 6S RNA homolog. *Nucleic Acids Res.* 33: 1949–1960.
- Zhaybayeva O, Swithers KS, Lapierre P, et al. (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. USA* 106: 5865–5870. <https://doi.org/10.1073/pnas.0901260106>.

## Gen- und Proteinsequenzen

### Aq\_880

ATGGATGTGTTCTCGACACGAGCGTTTTACAAAACCCGGAGATATACAGAACCTTTGAAGAGGATC  
AGAGAGGAGCTATGGAAACCTTTATACACCTTGCCCTGAATTCAGAGCTGAGTTTTACATGCCCACTTC  
CGTTTACACGGAATGAGAAAAATAATGGATGTGGGGAACTCTGGGCGGAATTTGAGATGGTGGAAG  
ATACGCTCTCCAGGAGGTTTCAATTAAGTGTCCCGCGGATTTTCTCTATGAGTTCATAGAAGAACTCA  
GATATAGGATTAATAAAGGTTTGAGGATAGCGGAAGAACACACGAGGGAAGCGAGCGGTTGTGAAGACGT  
CGGGAAATTAATAGCTAGGCTCAGGGAGAAGTACAGGGAAGCCCTCAGGCAGGGAATACTCGACAGTAAA  
GAGGACGTGGATGTGCTGCTCCTTGCTACGAACTGGACGGGGTTCTCGTTTCGGCGGATGAAGGCCTCA  
GAACATGGGCGGACAAGATAGGTATAAACTCATAGACCCTAAGAATTTCAAAAACATTTTAGAGAGCTT  
GGTAAGACACAGGTTTTAA

MDVFVLDTSVFTNPEIYRTFEEDQRGAMETFIHLALNSRAEFYMPSTSVYTEMRKIMDVGELWAEFEMVVK  
IRSPRRFQLTVPADFLYEFIEELRYRINKGLRIAEEHTREASGCEVDVGKLIARLREKYREALRQGILDSK  
EDVDVLLLAYELDGLVLSADEGLRTWADKIGIKLIDPKNFKNILESLVRHRF

### Aq\_707

ATGGAAATCCTTGTAATAGGCGGTGGGGTAATAGGACTATCCATCGGGGTAGAACTGCTCCTTGCGGGAC  
ACGATGTAAGACTCCTCGTCAGAGATAAAAGGGAAGGTGCTTCCTGGGTGGCGGGCGGTATGCTCGCCCC  
CTTTTCGGAAGGTCTGAGGGGAGATTTTTTAAGTTTTCCGTGGAAAGTCTAAGGCTCTGGCACGACTAC  
CTTGAAAGACTGAAGGAAATTCGGGGAGAAAAGTTTTCTTCTCTCAGGGAATACTCAGAGTAGCCTTTG  
ACGAGAAAGAAAAGAAATTGAGAGAGGACGTTAAAGTTTTACTCTCAGGAAGTGTGTCAAGACCTTAT  
CTCTTACGACGCTGAAGAACTCCTAAAAGAGTTTTCCCTACCTTTTCGAGTGAAGTCAGGTACGGAGTTATA  
TACGGAGATGAGGGAAATGTAGACACGGAAGAGCTTATGGAAGCCCTTTACGTTGCCTTTGAAAAACTCG  
GTGGAGAAATAATAAACGAGGACGCGGTAAAGATTCTGAGGAACGGAGAGAGAATAGAAAAAGTCTTCGG  
CTTTTTTGGAGGTTTACTGCGGATTACTACGTGTTTGCAACGGGTGCTTGGCTAAAAGAACACTTCGGC  
TTTCCCGTATTTCCCGGTAAAGGGGCAGATACTCAGGGTTGAAGCACCCCTCAAAGACTACGTTGTTTACT  
CCTCAAGGGCATAATAATCCCGAGGGAAAAGGACCTTTTAATAGGCGCGACAACCGAAAACGCTGGCTT  
TGACACGAAAACAACCTTTAGAAGGTGTTAAAAAACTTTCCGAAGGGGCGATATCGAGCGTCCCTCAGCTC  
AAGGAAGCTCAACTTCTTGAGGTAAGAGTCGGCTTTAGACCCGGAACGCCCGATGAACCTCCCGTTTTTT  
ACTTTGGAGAAAACCTTTTCAAGTTTACGGAGGGCATTACAGGAACGGCATACTCCTCGCTCCCGCAACCGC  
TGAGGTAGCCCTGAAGTTAATAGACCAGGGGGAGGTTTCGGAATACTTCAAGACTTTCTCCCTTACAGG  
TTCAAGGAGTAA

MEILVIGGGVIGLSIGVELLLAGHDVRLLRDRKREGASWVAGGMLAPFSEGLRGDFLRFSVESLRLWHDY  
LERLKEISGRKVFFSQGILRVAFDEKEEKKLREDVKFYSEQVCQNLISYDAEELLKEFPYLSSEVRYGVI  
YGDEGNVDTEELMEALYVAFKLGGEIINEDAVKILRNGERIEKVFGFFEEFTADYYVFATGAWLKEHFG  
FPVFPVKQILRVEAPLKDYVVYSSRAYIIPREKDLLIGATTENAGFDTKTTLEGVKKLSEGAISSVPQL  
KEAQLLEVRVGFRPGTPDELPVFYFGENFSVYGGHYRNGILLAPATAEVALKLIDQGEVSEYFKTFSPYR  
FKE

### Aq\_221

ATGAGTGTGGAGGAAAAAGCGAGGATAGGAACTCCGAAGAGCCTATTATTATAGAGACTGGAAAGTACG  
CAAACTCACCGACGGTTCCGTAGTAGTCAGGCAGGGCGGTACCGCGTTCTCGTGACCGCGGTTATGTC  
GGACGAGCCCATACCGACGTTGACTTTACCCCCCTGGCGGTAGACTACAGGGAAAGGGCTTCTGCGTAC  
GGTAGGATACCCGGAGGCTTACAAAAGAGGGAAGGAAAACCCACCGACAGGGAGATATTGGTCTCCAGAG  
TAATAGACAGACCAATAAGACCTCTCTTTCCCGAAGGATTTTTCCACGATGTTATAATAACCGCCCTCAC  
TTTATCCGCGGATGATAAGTACGACCCGGACGTTCTTGCAATAACCGGTGCTTCCGCAGCCCTCCACATA  
TCCAGAATTCCTTTCGAAGGTCCCATCGCGGGAGTCAGAGTCTGTAGGGTAAACGGGGAGTTTCGTGGCTA  
ACCCACCTACGAACAGAGGAAGGAAGCGGATCTCGACATAGTAATGGCGGGAACAAAAGACGCCATCGT  
AATGGTGGAAGGCGGAGGAAAGGAGATACCAGAAGAGGTTCTCGCAGACGCCCTGTCTCTCGGACTTGAC  
GCAATAAAGGAGTTATAGAAGCTCAGGAAAGACTGAGGGAAGAGGTAGGGAACCCCAAGTTTGAGTATC  
AGAAGGTTGAACTTCCCGAAGATATACTCAAGGCCCTTGAAGAAGAGTGACTCCAAAAACCTTGAGGC  
CTTTAACATAAAAGATAAAAAAGAAAGGTATTCAACCTTAGACAAGATAGTTGAGGAATTTATAGAAGCT  
CACCAAATCCCCGAAGAACTTCACTTTGCGGTTAAGTACTTCTACAAGAACTCGAAAGCAGGTTAATGC  
GTGAAAAGGTTCTAAAAGAAGGCGTAAGGATAGACGGAAGGAAACCAACGAGATAAGACCTATATGGAT  
AGAGGTACATCCCTTCGAAAGACCCACGGAATGCCATATTTACAAGAGGACAGACGAGGCTTACGTT

ACGGTAACCCCTCGGAACACCGGACGAAGCCCTGATAATAGAAACCATAGCGGAAGGGGAAGTCTTCAAGA  
GATTCATGCTCCACTACAGTATGCCTCCCTTTTCCGTTGGTGAGGCAAAACCCTGGGGTCCTCCGAGGAG  
AAGGGAAATCGGACACGGAGCCCTTGCGGAGAGGGCAATAGAACCATTACTTCCCTCCCGAAGAGGAGTAC  
CCCTTCATAATAAGGGTTGTGTCCGACATACTCGAGTCTAACGGCTCCACTTCCATGGCCACGGTTTGTG  
GTGCATCCTTGGCTCTCTTTGACGCCGGCGTTCCCATGAAGGACAACAAACATGTGGCCGGCATAGCCAT  
GGGACTTATCCTTGAAAAGGACAGATACGTTATACTCTCTGACATACTGGGGGACGAGGACCACCTCGGC  
GATATGGACTTCAAGGTTGCGGGAACATAAGACGGAATTACTTCCGTTCAAATGGACATAAAGGTAAAGG  
GAATTACAAAAGAGATAATGCTCGACGCTCTTAAACAGGCAAGGGAAGGAAGGCTCTACATACTTGAGAA  
GATGTACGAAGCTATACCCGAGCCGAGAAAAGAGCCTCACCCCTACACGCCGAAGGTTGAAGTTGTAGAG  
GTTCCCGAAGAAAAGCTCCCTCATAATAGGGCCGGGCGGTTCAACCGTTAAGAAGATTTACGACGAGA  
CGGGCGTTAAGGTATGGGTAGGTGAACAAGGAAAGGTTTACCTTTTTGTATTTCCCGCGGTGACGTGGA  
AAAAGCCAAACAAATGATTACAGGACATAGTAAGGGAAGTTGAGGTAGGTGCCGTTTACAAGGGAACCATA  
ACGAGGGTAGAACCGTACGGAGTGTTGCTTGAACCTCTGGCCAGGAAAGATAGGACTCCTCCACGTTTCCA  
AGATGGCAGAACCCGTAAGGAGCGCAACCGAAAAGTACAAAGTAGGAGAGGAAATAATAGTTAAGGTTCT  
AGACCTCGACGAACCTTGAAGACCGAGGTTTACAACAATAGGGATAGAAGACGTAGGTACCGAAAAGAAG  
GAGGTAAAACCTAAAGTAGGAGATGTATACGAAGGTAAAGTTGTTAGAGTAGAGCCATACGGTGCCTTCA  
TAGAGTACGCCCCCGGAAAGGTGGGACTCTTGCATGTTTCAAAGATGAAGGAAAGGGTAAAGGATGCAAG  
GCAAAAGTACAAGGTAGGAGACGTAGTAAAAGTCAAAGTTGTTGAAATAGACGAACAGGGAAGACCCAAG  
TTCACGGACGATGTCTAA

MSVEEKARIGNSEEP III IETGKYAKLTDGSVVVRQGGTAVLVTAVMSDEPIITDVDFTPLAVDYRERASAY  
GRIPGGFTKREGKPTDREILVSRVIDRPIRPLFPEGFFHDV IITALTLSADDKYDPDLAITGASAAHLI  
SRIPFEGPIAGVRVCRVNGEFVANPTYEQRKEADLDIVMAGTKDAIVMVEGGGKEIPEEV LADALFFGLD  
AIKEVIEAQERLREKVGKPKFEYQKVELPEDILKALEEECTPKILEAFNIKDKKERYSTLDKIVEEFIEA  
HQIPEELHFAVKYFYKKLESRLMREKVLKEGVRIDGRKPNEIRPIWIEVHPFERPHGNAIFTRGQTQAYV  
TVTLGTDEALI IETIAEGEVFKRFLHYSMPF SVGEAKPWGP RRREIGHGALAERAIEPLLPPEEY  
PF IIRVVS DILESNGSTSMATVCGASLALFDAGVPMKDNKHVAGIAMGLILEKDRYVILSDILGDEDLG  
DMDFKVAGTKDGITSVQMDIKVKGITKEIMLDALKQAREGRLYILEKMYEAIPEPRKEPHPYTPKVEVVD  
VPEEKAPLIIGPGGSTVKKIYDETGVKVWVGEQGVYLFVFPGGDVEKAKQMIQDIVREVEVGAVYKGTI  
TRVEPYGVFVELWPGKIGLLHVS KM AE PVR SATEKYKVGEI IIVKVLDLDELGRPRFTTIGIEDVGTEKK  
EVKPKVGDVYEGKVVRVEPYGAFIEYAPGKVGLLHVS KM KERVKDARQKYKVG DVVKVVEIDEQGRPK  
FTDDV

### Hhal\_2243 (in der Arbeit Hhal\_2443)

ATGCGCCGATTCTGCTCGATACCAGCGTCTTACCAACCCGGACGTCTACCTGCGCTTCGACGAAGAGC  
CGATGCAGGCCATCTCGGTCTTCTCGGGCTGGCAGCCGCGCCGACGCCGAGTTCTACATGCCCGGACC  
GGTCTATCAGGAACGTGCAACCTGCGCTCGATGGATCTGATCGGTGCCGAATTCGAGACCGAGGTCTAC  
ATCCGCTCGCCCCGCGCTTCTCCATGACCATCCCTCGGAGGTGCTCTACGAGTTCATCGAGGAGGTGC  
GCACCCGCATCCAGCGGGGCTGCGCATCGCCGAGGAACACGCCCGGCAGGCGGGTCAGGCGGAGTCGCT  
GCCGCCGAGCTGATACCCAGCTGCGCGAGCGCTACCGCGAGGCCATGCGCCGCGGCATCTCGATTCC  
CGCGAGGATATCGATGTGGTCTCTGCTCGCTACGAGCTGGACGCCACGCTGGTCTCCGCCGATGAGGGGA  
TGCGCAAGTTCGCCGAACGCATCGGCATCAAGCTGGTCAACCCGCGCTACCTGCGCGGGGTGATGCAGAA  
CCTGGCCGGCGACGACCCCGGTACGCGCCCCCTGCGGGCCGACGACCCCGCGGGTAG

MRRFVLDTSVFTNPDVYLRFDDEEPMQAI SVFLGLARRADAEFYMPGPVYQELCNLRSMDLIGAEFETEYV  
IRSPRRFSMTIPSEVLYEFIEEVTRIQRGLRIAEHARQAGQAESLPPELITQLRERYREAMRRGILDS  
REDIDVLLAYELDATLVSADDEGMRKFAERIGIKLVNPRYL RGV M QNL AGDDPGHAPPCGPDQ PAG

### Hhal\_1229

TTGTCGGCCGACGAGAGGTTCCCGCGGGGCGCACGGCTGCTGCGGCGGAGGGAATTCGAGGGGGTGCTTC  
GCGCCCCGGAATGGCGGACAGCCAACCGCTACTTCCGCGTTAGCGCACGGCAGAATCAGTGCGGGCGCCC  
GCGGCTGGGCCTGGCAGTCCCCAAGCGGGTGCTGCGTCTGGCCATACAGCGGCATCGAGTCAAGCGGATC  
ATCCGCGAGAGTTTCCGGGTCCGGCAGGGCGGCCTGCCGGACACGATTTTGTGCTCGGGGTGCGCGGGG  
CCGTGCGGAGGCGGACAACGCCACCCTATTAGTGCCCTGGAGGAACCTCTGGGCGCACGCGCGGGGGCG  
GCCATGCGATGGCTCCTGA

MSADERFPRGARLLRRREFEGVLRAP EWRTANRYFRVSARQNQCGRPRLGLAVPKRVLRLAIQRHRVKRI  
IRESFRVRQGGLPDHDFVVGVRGAVAEADNATLFSALEELWAHARGRPCDGS

**Thein\_2060**

ATGGCTTTTGAATTAGAAGTTGAAAGAGAAAGATTAATTCCTGACACCAGTGTTTTTACCAATCCTGATG  
TTTATCGCGAATTCGGAAATTCTCCCAATGAGGCTTTTCGCCAATTTTTTTAGAATTAGTGGCCAAAGCTCC  
CAATGTCAAAGTTTATATGCCTTCTTCGGTTTACGATGAACTAAAAAAATGCTCAAAAACCCAAAAATT  
CCCCCTAAAGCTAGAGCAGTATTTTCGAGTTAAATCACCTAAAAAATATGAAATTCAGATACCTGCTTTTT  
TACTTTACGAACTAATAGATGATATTAGGCAGCGTATTAATAAAGGTTTAAGGGTGGCTGAAGAGGCAGT  
AAGGGCCTCTGCGCACAAAAGTCCTGACGAAATATTAATGCCCTCAGACGTAAATACCGCGAGGCCTTA  
CGAGAAGGTATTGTTGATTCCAAAGAAGACATAGAATTAATTCTTTTAGCTCTTGAACCTGATGGGTAT  
TACTTTCTGCTGATCGCGGTATTCTTAACCTAGCAGATAAAATGGGTATAAGGTGGATACCACCAGAAGA  
AATAAGAGATACCATAGAAGGCCTTATTTACAGTCCTTAA

MAFELEVERERLIPDTSVFTNPDVYREFGNSPNEAFANFLELVAKAPNVKVMPSSSVYDELKKMLKNPKI  
PPKARAVFRVKSPPKYEIQIPAFLLYELIDDIRQRINKGLRVAEEAVRASAHKSPDEILNALRRKYREAL  
REGIVDSKEDIELILLALELDGLLLSADRGILNLADKMGIRWIPPEEIRDITIEGLIYSP

**Thein\_0634**

TCAACCCAGTCGTATCAAGCGTAAAAGATGCCATGGTTTTTCGAGCTCGCATGAGAACTCGTTCAGGAAGG  
CGTATTTTGGCTAACAGGCGGCGTAAGGGACGTTGGCGGCTTACGGTATAAAAAAGAATCTTTTACCAA  
AGATGAGCGTCTAAGGCGTCGCAGGGAGTACGAAAGGGTTTACCGTCAAGGTAAACGCCTTTCGTTACCT  
TATTTAAGAATAACTAGCCCCAAATCAATTAGGCCATTCAAGATTAGGTCTTAGTGTTTCCAAAAAGG  
TAGGAAATGCGGTGAAACGTAATTATATTAACGAATTTTGCAGAGAAGTATTTAGAAGAAATAGAGATAT  
CTTCCCTTTAAGTCATGATATAGTTATAATTCCTAGGGCTAAAATATTAGAAATCCCTCAAGATAAAATT  
AAAAAAGATTTAGTTGAATTATTTAAAAAGTATGAGAAAAATAGTTATAAAAGTTATTTAAATATATCAA  
TTATTACTTTACCTTTTTTGCCTCAAAGTTGTCGTTTTTATCCTACTTGTAGTCATTATGCTATAGAAT  
CTATTTCTAAGTTTGGGGT

MAAYGIKKESFTKDERLRRRREYERVYRQ GKRLSLPYLR IILAPNQLGHSRLGLSVSKKVGNVAVKRN IYK  
RILREVFRRNRDIFPLSHDIV IIPRAKILEIPQDKIKKDLVELFKKYEKNSYKSY

**Tvnir\_2893**

ATGCGTCGCTTCGTTCTGGATACCAGCGTTTTTACCAATCCGCACGTGTTTCGGGCAGTTTGGCGAGGATC  
CCGGGGATGGCCTTCGCACCTTCCTGCGTCTCGCCAGCGCTGCCCGGCCGAGTTCTTCATGCCGCTGTC  
GGTCTACGACGAGTTCGGACGATGCGCGATCTGGCTGAACTCGCGGCCGACTTCGAGACGGTCGTCTGG  
GTGCGCTCGCCAGGCGTTTACGCTGACGATCCCCAGCGACATCCTGTACGAGTTCATCGACGAGGTTTC  
GCACCCGGATCGATCGCGGTCTGCGCATCGCGGAGGAACATACCAAGCGCGCCGGCGCGGCCGCGGATAT  
GCCCCCGGAGATCATCACCCAGCTGCGCGAGCGGTACCGGGAGGCAATGCGCAAGGGACTCGTGGATTCC  
CGGGAAGACGTCGATGCAGTGCTGCTGGCGCTGGAGCTCGACGCCGAAGTGGCCAGCGCGGACGAGGGCA  
TGCGCAAGCTCGGCAACCGCATGGGCATCAAGCTGGTCACCGCCGACTACCTGCGCCGGGTCTATGGAGAA  
CCTAGCGGAATCTCAGGGCGGAGGGGTTCCTCAAGACGTCCGGCTGA

MRRFVLDTSVFTNPHVFGQFGEDPGDGLRTFLRLAQRCPAEFFMPLSVYDEFRTMRDLAELAADFETVVW  
VRSPRRFTLTIPSDILYEFIDEVRTRIDRGLRIAEEHTKRAGAAADMPPEIITQLRERYREAMRKGLVDS  
REDVDAVLLALELDAELASADEGMRKLG NRMGIKLV TADYLRRVMENLAESQGGGVPKTSG

**Tvnir\_3809**

CTCGCTCCGGTCGCCGGGTGATCAACGCCCCGCCGCGCCAAGGGCCGCAAGCGCCTCTGCCCCGTAAGCAGC  
GCGTCTCTGTTGTGCCCGCCCAGGCGTGCACGCCGTGAGCGGGCCAGCTCCGACGCCGGGGACCGTTTT  
CCCCGTAGCCTGCGCCTGCTCGACGGCCGCGCGTTCCAGCGCGTGTGTGCCCGCGCCCGCCGCATCCCCG  
GCCGGCGTTTCATGCTGTTGTACCGGGTTTGCGCACCGCCGCTCGATCATCCCCGGCTCGGGCTTGCGAT  
CTCGAAGAAGCACGCCCGGCGCGCGGTGGACCGCAACCGGATCAAGCGCGTTCGCCCGGGAGGCGTTTCGC  
GCGCGCGTCCGTTACTGATCGCGGTGGACATCGTCTGCTGTCCCGCGCCGGGATCGCCGATGCAAGGCT  
CCCGGGAATTGCGCGCGGAACTCGACCAGTTGCTGGAGCACGTGAAATGATCCGGACCCTGATCCTGATG  
CCGATCCGGCTGTACCAGTACGCGATCTCGCCCTTCATGGCGCCGCATTGCCGGTATTATCCCAGCTGCT  
CCGAGTA

MSGPAPT PGDRFPRSLRLLDGRAFQRVFARARRIPGRRFMLLYRVCAPPLDHPRLGLAISKKHARRAVDR  
NRIKRVAREAFRRRRSLIAVDIVVLSRAGIADAGSRELRAELDQLLEHVK



**Mlg\_2556**

ATGCGCCGATTTCGTGCTGGACACCAGCGTTTTACCAACCCCCACTGCGCCGTGCAGTTCGGCGAGGACC  
CGCTGGCCGGGGTGCAGACCTTCTGACCTGGCCCCGGCGCTGTCCAGCGGAGTTCTACATGCCGCTGTC  
GGTCTATGACGAGTTCCGGCGCATGCGGGACCTGGCGGAGATGGCGGCGGACTTCGAGACCGATGTCTGG  
GTCCGTTACACGCGCCGGTTTTCAGCATGACCATCCCGGCGGAGATCCTCTACGAGTTCATCCACGAGCTGC  
GCGGCCGCATCGACCGCGGGCTGCGGATCGCCGAGGAGCACACTCGCCAGGCGGGGGCGGCCACCGACAT  
GCGTCCGGAGCTGATCGCCAGCCTGCGCGAGCGCTACCGGGAGGCCATGCGCAAGGGGGCTGGTGGACTCC  
CGCGAGGACGTGGACGCGGTGCTGCTGGCCATGGAGCTGGACGCCGAGCTGGTCAGCGCCGACGAGGGCA  
TGCGCAAGTTGGGCAATCGGATGGGCGTGAAGCTGCTCACC GCGGACTATCTGCGCCAGGTGATGGAGAA  
CCTGGGTGCCGGCCATTAG

MRRFVLDTSVFTNPHCAVQFGEDPLAGVQTFLLHLARRCPAEFYMPLSVYDEFRRMRDLAEMAADFETDVW  
VRSPRRFSMTIPAEILYEFIHLELRGRIDRGLRIAEHTRQAGAATDMRPELIASLRERYREAMRKGLVDS  
REDVDAVLLAMELDAELVSADEGMRLGNRMGVKLLTADYLRQVMENLGAGH

**Mlg\_RS14575**

AGCGCACCCACGGCTTCCGGGCCCCGTATGGCCACCAAGAACGGGCGTGCCGTGCTGGCGCGGCGTCGCGC  
CAAGGGTCGCCATCGTCTCTGTCCCTGATTTGGTGGCGGCCCGAGGCCATGGCCCCTGCGCCTTCCCGCG  
CCGTGCGCGACTAACGCGTCCCCAGGACTACCGGCGCGTCTTTTCCGGCGCCCAGAAGGCCGCCGATCGC  
TACTTCACCGTGTTGGCAGCGGCAACGACAGGTTCGGCCCCGCGTCTGGGCCTGGCCATCTCCCGCAGGG  
TCGCGCCTCGTGCCGTGGACCGGAACCGCCTCAAGCGGATCGCCCGGGAGGTCTTCCGGCAACGCCAGGC  
GGCGCTGGGACAGCGAGACTACGTGGTCCTGGCGCGGCCGGCGGCGAGGACGGCCAGTAAGGCCCGGCTC  
CACCGCAGCCTGGGGCGGCTCTACGATCGGCTGAACGACGGCAACAGGCACTAAGGGGCGGTTTCGCCCTC  
CCCGCCTTTCCGGTACCCATCGGTGTCCCCCTACCCAACCCCCCTGGTGAGCGGTTCATGGAAAACGAGCGC  
ATATTCCTCTTTC

VAARGHGPCAFFRRRLTRPQDYRRVFSGAQKAADRYFTVLARGNDQVGPRLGLAISRRVAPRAVDNRNL  
KRIAREVFRQRQAALGQRDYVVLARPAARTASKARLHRSRLGRLYDRLNDGNRH

**Mth\_243**

ATGCTTGCAAAGCAGAGATTTCGTACTTGATACAACGGCATTTCACAGATAACCAGCTGCGAGAACTGCTGG  
GTGATGGCGACCTCAACCTCTCGGTTGACAGGATGCTTGACCTCATAGCGAGAAGCAGGATAAACTCAA  
CATAAGCTGTACATGCCTCCCATCACCTACAAGGAGTTCACAGATTACATGACACGCTACGACTGCCCG  
GAGGAGACCCTTATAAAGGCCGAGACATGGATAGTTAAGAAGACACCCAACCGCTACGACACACAGATAC  
CCTCCCAGATATTCTATGAATACGTCCATGATATACGGGAGAGGATGAACAAGGGCCTCAGGATCTCAGA  
GACCCTGCTCTGGGAGGCAGGCATACAGTCCATCATAATGGCCTCCCGTGACGTCAAGAAGACAGAGATC  
GAGAGCGAGGTGCTTGAAAGGCCATAAAGGACCTCAGGAAGAAATACAGATCAGCCCTCCGCAAGGGAA  
CCCTTGACAGTGCCCCGACCTTGATGTGCTGCTCCTTGCCAAGGAAGTGGGGGCGAGGGGTCGTGGCGGC  
CGATGATGGTATAAGGGTCTGGGCTGAGCGACTGGGTCTGAGGTTCTCAATGCCACATCCTTCCCCAAG  
ATGCTCAAGGAATACCTTAAATATTATGAATGA

MLAKQRFVLDTTAFSTDNQLRELLGDGDLNLSVDRMLDLIARSRIKLNISCHMPPITYKEFTDYMTRYDCP  
EETLIKAETWIVKKTNPNYDTQIPSQIFYEYVHDIRERMNKGRLRISETLLWEAGIQSIIMASRDVKKTEI  
ESEVLGKAIKDLRKKYRSALRKGTLD SAPDLVLLLAKELGAGVVAADDGIRVWAERLGLRFLNATSFPK  
MLKEYLKYYE

**Mtbma\_c06930**

ATGCTTGCCAAGCAGAGATTTGTCCTTGATACAACGGCCTTTACAGATAACCAGCTCCGGGAGATGCTGG  
GTGACGGTGACCTCAACAGGTCACTGGATACCATGCTCGACATCATCGCAAGGAGCAGGATAAAGCTCAA  
CATAAGCTGTACATGCCCCCATAACCTACAAGGAATTCCTGACTACATAACACGGTATGAATGCCCC  
GAGGAGACCTTGGTGAAGGCCGAGACATGGATAGTCAAGAAGACACCCAACCGCTACGACACCCAGATAC  
CCTCCCAGATATTCTATGAATACGTCCATGATATAAGGGAGAGGATGAACAAGGGGCTCCGGATTTCAGA  
GACCCTCCTCTGGGAGGCCGGGATACAGTCCATAATCATGGCATCCCGTGGTGTTAATAAGATTGAAATC  
GAGGGTGAAGTGCTCGGGAAGGCCATAAAGGACCTCAGGAAGAAGTACCGTTCAGCCCTCAGGAAGGGGA  
CCCTTGACAGTGCCCCGACCTCGACGTCTCCTTCTTGCCAAGGAGCTCGGGGCGAGGTGTTGTGGCTGC  
TGATGACGGTATAAGGGTCTGGGCGAGCGGCTGGGACTTAGGTTCTGAATGCCACATCCTTCCCCAAG  
ATGCTCAAGGAATACCTTAAATATTATGAATGA

MLAKQRFVLDTTAFSTDNQLREMLGDGDLNRSVDTMLDIARSRIKLNISCHMPPITYKEFTDYITRYECP  
EETLVKAETWIVKKTNPNYDTQIPSQIFYEYVHDIRERMNKGRLRISETLLWEAGIQSIIMASRGVNKIEI

EGEVLGKAIKDLRKKYRSALRKGTLD SAPDL DVLL LAKELGAGVVAADDGIRVWAERLGLRFLNATSFPK  
MLKEYLKYYE

### Pfc\_09185

ATGAGGGCGGGTGTACGGGCGGGCCACAGATAATAACGAAGGTGATTACTATGATTAAATTTGTTCTAG  
ATACCAGCATATTCGTAAATCCAGATGTGAGGAAAAAATTTGGAGAAACACCAACAGAAGCCATGAAGAC  
CTTCCTCCATTACGCTGAAAACCTGTTTGGAAAAGTTGAATTTTATATGCCTCCTGGGATATATCGCGAG  
CTTATGCATTTTCGTTGAGGAGGAAGAGGTATCACCAGACATAGAGTTGTACATAATAAAAAAGCCTCCCA  
ACGTTACGACATAAAAAATACCTGCATTTGTAGTTTATGAGCTTATAGAGGATATAAGGAGGAGGGTAGA  
TAAAGGGCTTAGGGTTGCAGAAAAGGCCGTGAGGGAGAGCGTTATAGACACGAGTAATGTGGACAAGATA  
ATTCAAAACTTAGAAGAAATTATAGAAAAGCTTTGAGAGAGGGTATTCTTGACAGTAAAGAAGACTTCG  
AGCTTATACTACTCGCGAAAGAGATCGATGGGATAATAGTCTCCGCAGATGTTGGGATATTAACATGGGC  
GGAGAAGATGGGGATAAAGTGGGTGATGCTTTTAAATTTAAAGAAGTTCTTAGCGAGCTAGTTGAGAAG  
TTTAAGAGAAGTGAAAGTGAAAAGGAAAGGAAATGA

MRAGVTGGPQIIITKVITMIKFVLDTSIFVNPDVRKKFGETPTEAMKTFLHYAENLFGKVEFYMPPIYRE  
LMHFVEEEEVSPDIELYIIKKPPNVHDIKIPAFVVYELIEDIRRRVDKGLRVAEKAVRESVIDTSNVDKI  
IQKLRRNYRKALREGILDSKEDFELILLAKEIDGIIVSADVGILTWAEKMGIKWVDAFKFKEVLSSELVEK  
FKRSESEKERK

### Minf\_0007

GTGAGAAAGTTTGTTTTAGACACCAGCATATTTACCAATCCGGATGTATCCCAGCAATTCGGAGATGATG  
ACAGTTCCGTGTTTTCCGGCTTTCTATCTTTAGCCTCCAAGGCTGAAGCCCTTTTTTACATGCCCTTGTC  
CGCCTACAAGGAGCTCAAGCTCATCAAGGGTACGGCTTTTTTTTCTTCTGACTTTGAATTCATCGTCCAT  
ATTCGTTCTCCAAGAAAGTTCAATCTTCTTATACCCGGTTCCGTTCTTTATGATTTTATCGATGAAATAA  
GGAAAAGAATTGACCATGGGCTGAGAATTGCAGAAGAGCACTTAAACTTGCCGGTCCAGACGGTGGAAC  
TCCCATTTCCCTTTGATCAACTGGTTAATCGATTAAGAAATCGATACCGCGAAGCCTTAAGGCAGGGTATT  
CTGGATAGCAGGGAAGACGTTGACGTTATCCTCCTTGCCATGAAGTAGAGGCAGTTCTCGTTACCGCCG  
ATGAGGGTATGAAAAAATGGGCTGATAAAATGGGAATAGAAATTTTAAATGCAAGTTATTTGAAAGACTG  
CTTAAAAAACCTAATCAAACCTGAGGGCACCGGAGAGCCAGTTTCTTAA

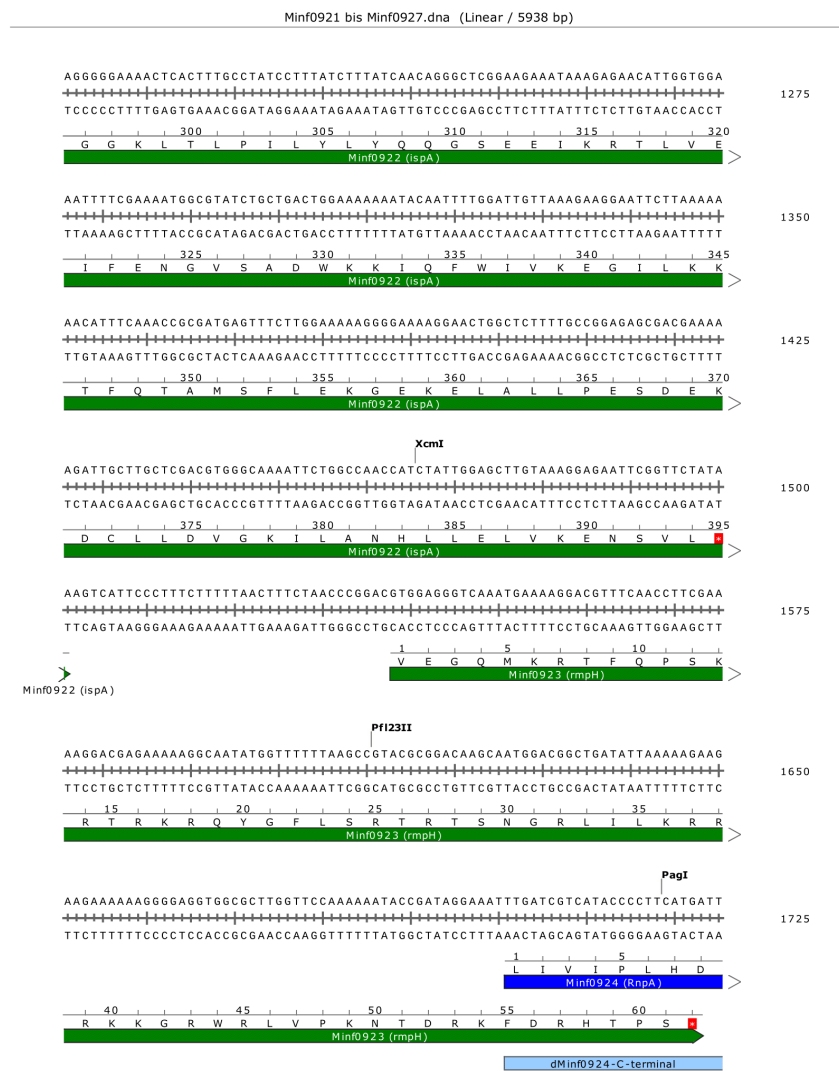
MRKFVLDTSIFTNPVDSQQFGDDSSVFSGLSLASKAEALFYMPLSAYKELKLIKGTAFFSSDFEFIVH  
IRSPRKFNLLIPGSVLYDFIDEIRKRIDHGLRIAEEHLKLAGPDGGTPIPFQDLVNRLRNRYREALRQGI  
LDSREDVDVILLAYELEAVLVTADEGMKKWADKMGIEILNASYLDCLKNLIKPEGTGEPVS

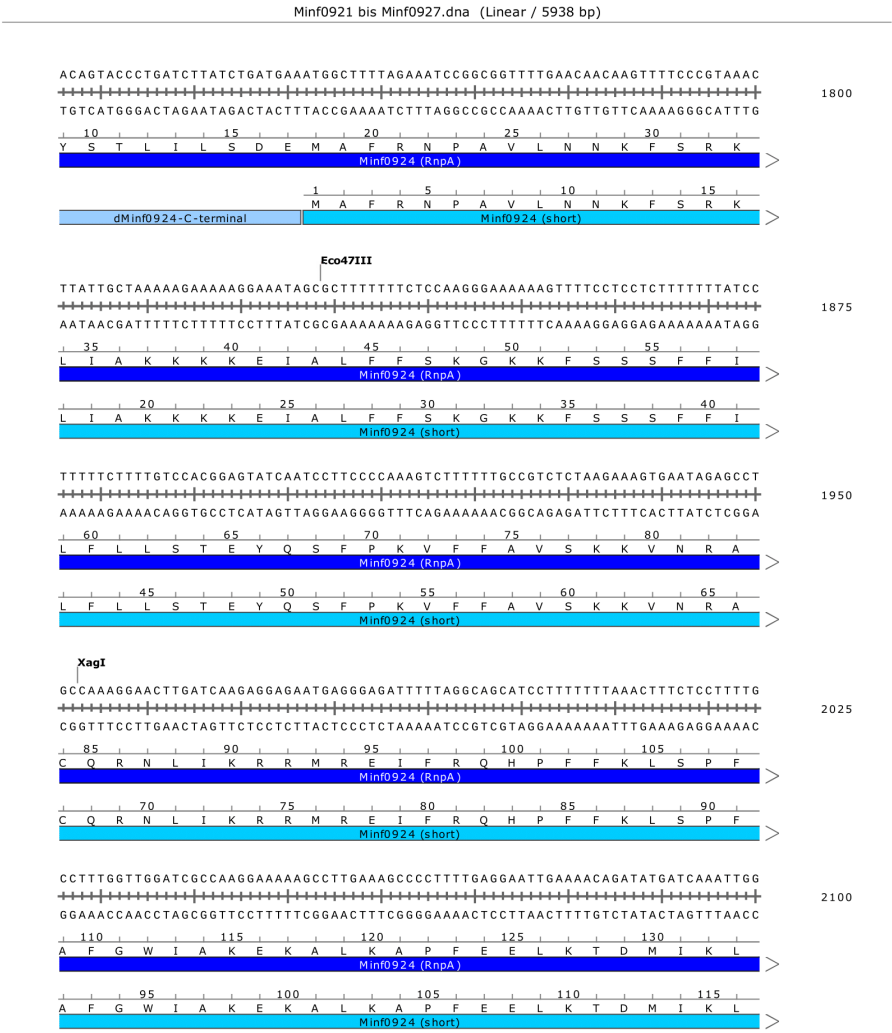
### Minf\_0924

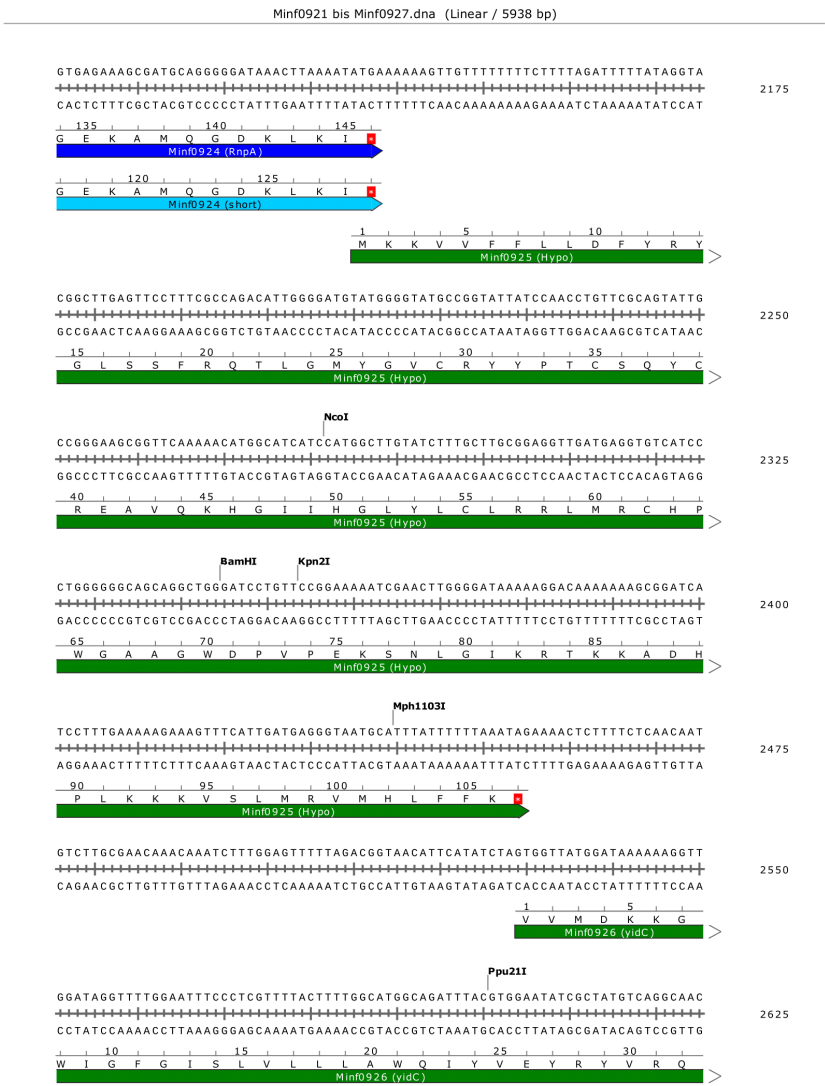
TTGATCGTCATACCCCTTCATGATTACAGTACCCTGATCTTATCTGATGAAATGGCTTTTAGAAATCCGG  
CGGTTTTGAACAACAAGTTTTCCCGTAAACTTATTGCTAAAAAGAAAAAGGAAATAGCGCTTTTTTTCTC  
CAAGGGAAAAAAGTTTTCTCCTCTTTTTTTTATCCTTTTTCTTTGTCCACGGAGTATCAATCCTTCCCC  
AAAGTCTTTTTTGCCGTCTCTAAGAAAGTGAATAGAGCCTGCCAAAGGAACTTGATCAAGAGGAGAATGA  
GGGAGATTTTTAGGCAGCATCCTTTTTTTAACTTTCTCCTTTTGCTTTGGTTGGATCGCCAAGGAAAA  
AGCCTTGAAAGCCCCCTTTTGAGGAATTGAAAACAGATATGATCAAATTGGGTGAGAAAGCGATGCAGGGG  
GATAAACTTAAATATGA

MIVIPLHDYSTLILSDEMAFRNPAVLNNKFSRKLIAKKKKEIALFFSKGKKFSSSFFILFLLSTEYQSFP  
KVFFAVSKKVNACQRNLIKRRMREIFRQHPFFKLSPFAGWIAKEKALKAPFEELKTDMIKLGEKAMQG  
DKLKI

## NCBI- Annotation von Minf0924







## Publikationen

### Veröffentlichungen in Fachzeitschriften

*Proceedings of the National Academy of Sciences of the United States of America (PNAS)*

**Minimal and RNA-free RNase P in *Aquifex aeolicus***

Astrid I. Nickel, **Nadine B. Wäber**, Markus Gößringer, Marcus Lechner, Uwe Linne, Ursula Toth, Walter Rossmanith, Roland K. Hartmann

**Published: 11.09.2017**

[www.pnas.org/cgi/doi/10.1073/pnas.1707862114](http://www.pnas.org/cgi/doi/10.1073/pnas.1707862114)

*International Union of Biochemistry and Molecular Biology (IUBMB)*

**Homologs of *Aquifex aeolicus* Protein-Only RNase P Are Not the Major RNase P Activities in the Archaea *Haloferax volcanii* and *Methanosarcina mazei***

Thandi Schwarz\*, **Nadine B. Wäber\***, Rebecca Feyh, Katrin Weidenbach, Ruth Schmitz-Streit, Anita Marchfelder, Roland K. Hartmann

**Published: 08.07.2019**

<https://doi.org/10.1002/iub.2122>

### Veröffentlichungen in Büchern

*Encyclopedia of Microbiology (4th Edition)*

**Aquificae**

**Nadine B. Wäber** and Roland K. Hartmann

**Published: 09.08.2019**

**Book ISBN: 9780128117361**

### Veröffentlichungen auf Fachkonferenzen

**GbM Tagungen (Bonn, Deutschland)**

**Comparative tRNA processing in *Aquifex aeolicus* and *Bacillus subtilis*: substrate analysis**

Nickel, A.I., **Waeber, N.B.**, Gößringer, M, Helmecke, D.J, Walczyk, D, Hartmann, R.K

2014 - Poster

**AaRP and HARPs – Novel protein-only RNase Ps exclusively found in Prokarya**

**Wäber, Nadine B.**; Nickel, Astrid I.; Gößringer, Markus; Lechner, Marcus; Schwarz, Thandi; Weidenbach, Katrin, Schmitz-Streit, Ruth; Marchfelder, Anita; Rossmanith, Walter; Hartmann Roland K.

2018 - Poster

**IRTG (Gießen und Marburg, Deutschland)****RNase P in *Aquifex aeolicus*- last mystery of tRNA processing****Waeber, Nadine B.**; Nickel Astrid; Hartmann Roland K.

2015 – Poster

**The last mystery of tRNA processing**Identification and characterization of the RNase P of the hyperthermophilic bacterium *Aquifex aeolicus***Nadine B. Wäber**

2015 - Vortrag

**The last mystery of tRNA processing**Identification and characterization of the RNase P of the hyperthermophilic bacterium *Aquifex aeolicus***Wäber, Nadine B.**; Nickel Astrid I.; Gößringer, Markus; Helmecke, Dominik J.; Hartmann, Roland K.

2015 - Poster

**Tag der Chemie (Marburg, Deutschland)****The last mystery of tRNA processing**Identifizierung und Charakterisierung der RNase P des hyperthermophilen Bakteriums *Aquifex aeolicus***Nadine B. Wäber**

2015 - Vortrag

**VAAM (Marburg, Deutschland)****Comparative study of tRNA processing in *Aquifex aeolicus* and *Bacillus subtilis***Nickel, A.I.; **Waeber, N.B.**; Goessringer, M.; Helmecke, D.J.; Walczyk, D.; Hartmann, R. K.

2015 - Poster

**tRNA Conference (Kyllini, Griechenland)****Proteinaceous vs. bacterial RNase P: a comparative mechanistic analysis****Nadine Wäber**, Dennis Walczyk, Sören Seidler, Markus Gößringer, Roland K. Hartmann

2014 – Poster

**tRNA Conference (Jeju, Süd-Korea)*****Aquifex aeolicus* uses a novel RNase P***In vitro* and *in vivo* characterization of the catalytic domain**Nadine B. Wäber**; Astrid I. Nickel; Markus Gößringer; Marcus Lechner; Walter Rossmanith; Roland K. Hartmann



2016 – *Poster*

**tRNA Conference (Straßburg, Frankreich)**

**Novel protein-only RNase Ps in prokarya**

comparison of AaRP and archaeal HARPs

**Wäber, Nadine B.**; Nickel, Astrid I.; Gößringer, Markus; Lechner, Marcus; Schwarz, Thandi; Weidenbach, Katrin; Schmitz-Streit, Ruth; Marchfelder, Anita; Rossmann, Walter; Hartmann Roland K.

2018 - *Poster*

**Small non-coding RNAs (Hirschegg, Österreich)**

**Another way of tRNA processing in bacteria**

*Aquifex aeolicus* uses a novel type of RNase P

**Nadine B. Wäber**

2016 – *Vortrag*

**AaRP and its family**

**A novel class of protein-only RNase P in prokaryotes**

**Nadine B. Wäber**

2017 – *Vortrag*

**Protein-only RNase P in prokaryotes**

**Evolution and structural characterization of AaRP and HARPs**

**Nadine B. Wäber**

2018 – *Vortrag*

**AaRP and HARPs**

**Identification, characterization and substrate recognition**

**Nadine B. Wäber**

2019 – *Vortrag*

**Lebenslauf**

Nadine-Bianca Wäber, geb. 18.02.1988 in Lich (Deutschland)

**Schulische Ausbildung**

1994-1998	Grundschule Langsdorf Langsdorf, Deutschland
1998-2004	Dietrich-Bonhoeffer Schule Lich, Deutschland
2004-2007	Herderschule Gießen, Deutschland

**Abitur****Akademische Ausbildung**

2007-2010	Biologie Studium B.Sc Philipps-Universität Marburg, Deutschland FB Biologie
2010	<b><u>Bachelorarbeit</u></b> am Max-Planck-Institut für terrestrische Mikrobiologie AG Sonja Verena Albers <b>Auswirkungen einer pyrE Deletion auf die Biofilmformation in <i>Sulfolobus acidocaldarius</i></b>
2011-2013	Biologie Studium M.Sc Karlsruher Institut für Technologie Fakultät für Chemie und Biowissenschaften
2013	<b><u>Masterarbeit</u></b> am Institut für funktionelle Grenzflächen AG Jörg Overhage <b>Einfluss zyklischer di-Nukleotide auf die Signaltransduktion Gram-negativer Bakterien</b>
ab 2014	<b>Dissertation</b> Philipps-Universität Marburg FB Pharmazie  Doktorarbeit am Institut für pharmazeutische Chemie AG Roland K. Hartmann

## Danksagung

Zunächst möchte ich mich bei meinem Betreuer und Erstgutachter **Prof. Dr. Roland Hartmann** bedanken, der mir ermöglichte dieses umfangreiche und interessante Projekt in seinem Arbeitskreis anzufertigen, sowie auch für die anregenden Diskussionen, Besprechungen und Verständnis während dieser Zeit.

**Prof. Dr. Peter Friedhoff** möchte ich für die Übernahme des Zweitgutachtens danken.

**Prof. Dr. Klaus Reuter** und **Prof. Dr. Moritz Bünemann** danke ich für die Teilnahme an meiner Disputation.

Den Mitgliedern des International Research Training Group GRK 1384 möchte ich danken für die Möglichkeit an (inter)nationalen Konferenzen teilnehmen zu können, sowie die Möglichkeit mich weiter zu bilden.

Ein weiterer Dank geht an alle Kooperationspartner, welche mir ermöglichten, weiterführende Studien an unseren AaRP/HARP durchzuführen und uns mit ihren Resultaten unterstützten.

Nun zu den lokalen Danksagungen. Zunächst möchte ich mich **Markus** für Unterstützung, Betreuung und Diskussionen während dieser Zeit bedanken. **Marcus** danke ich ebenfalls für die Unterstützung, Betreuung und Diskussionen, sowie den Linux-Support. Meine Fähigkeiten Linux zu zerstören haben ihn öfters verzweifeln, aber nie aufgeben lassen. Meine Büromädels **Isabell, Becki, und Jana**, sowie unserem Hahn im Korb **Dominik** danke ich für die lustigen Stunden, die konstruktiven Besprechungen und natürlich für die Unterstützung in allen Bereichen. Allen anderen Mitgliedern der Arbeitsgruppe (**Jana, Laura, Marietta, Wiebke, Sweetha, Amri, Simone, Paul, Arnold und Kerstin**) danke ich ebenso für die lustigen Stunden, besonders beim Kochen, Privat als auch auf der Arbeit.

Allen ehemaligen Kollegen danke ich für den guten Einstieg in das Promotionsleben.

All meinen Freunden und Bekannten außerhalb des Labors danke ich für die Jahre lange Freundschaft und den Freizeitausgleich. Vor allem meinen Sport-Kumpanen, sowie meiner Kneipencrew und natürlich den Gästen, die immer versucht haben zu verstehen, was ich da eigentlich mache.

**Ein letzter Dank gilt meiner Familie und meinen Freund, welche mir durch alle Höhen und Tiefen in dieser Zeit zur Seite standen und mich nie aufgegeben haben.**

## Erklärung

Ich versichere, dass ich meine Dissertation

**Struktur- und Funktionsanalyse der  
Protein-basierten RNase P des hyperthermophilen Bakteriums  
*Aquifex aeolicus* und homologer Enzyme**

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den.....

.....

(Nadine Bianca Wäber)